Identification and Classification of Helicopters using Neural Nets

by

Sohail Akhtar

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

SYSTEMS ENGINEERING

December, 1998

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

IDENTIFICATION AND CLASSIFICATION OF HELICOPTERS USING NEURAL NETS

BY

SOHAIL AKHTAR

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

SYSTEMS ENGINEERING

December, 1998

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

This thesis, written by

Sohail Akhtar

under the direction of his Thesis Advisor, and approved by his Thesis committee, has been presented to and accepted by the Dean, College of Graduate Studies, in partial fulfillment of the requirements for the degree of

Master of Science in Systems Engineering

Thesis Committee :

Dr. M. El-Shafei Ahmed, **Chairman**

Dr. M. Shahgir Ahmed, **Co-Chairman**

Dr. Lahouari Cheded, **Member**

Dr. A. A. Andijani,
Chairman, Dept. of Systems Engineering

Dr. A. M. Al-Shehri
Dean, College of Graduate Studies

Date : 14/12/99

IDENTIFICATION AND CLASSIFICATION OF HELICOPTERS USING NEURAL NETS

SOHAIL AKHTAR

SYSTEMS ENGINEERING

December, 1998

Dedicated To

My Family Members
Whose Prayers, Inspiration And Elegant Support
Led To This Accomplishment

Acknowledgments

All praise be to Allah Subhanahu wa Ta'ala, the Lord of the Worlds, for having guided me at every stage of my life. I seek His mercy, favor and forgiveness. I feel privileged to glorify His name in the sincerest way through this small accomplishment.

I wish to express my sincere gratitude and thanks to my thesis advisor Dr. M. El-Shafei Ahmed. His constant encouragement and constructive criticism throughout this research work is highly appreciated. I would also like to extended my appreciation to the co-advisor Dr. M. Shahgir Ahmed. His comments, suggestions, guidance, and critique made the work possible. I would like to extend my sincere and special thanks to my committee member Dr. L. Cheded for his constant help, guidance, support and valuable suggestions throughout this work.

I am also indebted to the department chairman, Dr. Abdul-Basit Andijani and other faculty and staff members for their support. I would like to extend my sincere thanks to all my friends at KFUPM who made my stay at this university memorable.

Last, but not the least, thanks are due to the members of my family for their emotional and moral support, patience, encouragement and prayers throughout my academic career.

Contents

List Of Figures

List of Tables

Abbreviations

ANNs	Artificial neural networks
DFT	Discrete Fourier transform
FIR	Finite impulse response
HMM	Hidden Markov models
LP	Liner prediction
LPC	Linear predictive coding
LSF	Line spectral frequencies
LSP	Line spectral pairs
PARCOR	Partial autocorrelation
PSD	Power spectral density
SNR	Signal-to-noise ratio
STFT	Short-term Fourier Transform
WT	Wavelet transform

Abstract

Name: **Sohail Akhtar**

Title: **Identification and Classification of Helicopters Using Neural Nets**

Major Field: **Systems Engineering**

Date of Degree: **December 1998**

Artificial Neural Networks (ANNs), in combination with time series analysis techniques, are applied for the recognition of helicopter sound. Coding techniques based on linear prediction coefficients (LPC) have been applied to obtain spectral estimates of the acoustic signal. Feature vectors consisting of these spectral estimates are used for the training and testing of the ANN classifiers. Other forms of the LPC parameters such as reflection coefficients, cepstrum coefficients and line spectral pairs have also been used as feature vectors for the training and testing of the ANN classifiers. We have also investigated the use of wavelet transform for signal de-noising prior to feature extraction. The performance of various feature extraction techniques is evaluated in terms of their detection and classification accuracy. The proposed helicopter recognition system consists of two stages; the first stage detects the presence of a helicopter and the second determines the type of the helicopter.

MASTER OF SCIENCE DEGREE

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS
Dhahran, Saudi Arabia
December, 1998

خلاصة الرسالة

الاسم : سهيل اخستر

العنوان : تمييز وتصنيف نوع الطائرات العمودية باستعمال الشبكات العصبية

التخصص : هندسة النظم

تاريخ الشهادة : ديسمبر ١٩٩٨

تم في هذا البحث استعمال الشبكات العصبية الصناعية مع طرق لتحليل المتواليات الزمنية للتعرف على البصمة الصوتية للطائرات العمودية. وكذلك استعملت طرق لتشفير الإشارات مبنية على معاملات التنبؤ الخطي لتقدير طيف الإشارات الصوتية. وتم بناء متجهات للملامح من هذه الخصائص الطيفية واستعملت لتدريب واختبار مصنف يعمل بالشبكات العصبية الصناعية.

كما استعملت ملامح أخرى مستخلصة من معاملات التنبؤ الخطي مثل معاملات الانعكاس ومعاملات سبسترم ، وخطوط الطيف الزوجية في تدريب واختبار نظام الكشف والتعرف على الطائرات العمودية. وقد درسنا أيضاً إمكانية استعمال المويجات لإزالة الضوضاء وتأثير ذلك على تحسين القدرة على التمييز. ويتكون النظام المقترح من مرحلتين ، في المرحلة الأولى يتم الكشف عن وجود الطائرات العمودية. وفي الثانية يتم تمييز نوع الطائرة العمودية. وتم تقييم الطرق المختلفة للتصنيف طبقاً لدقة تمييزها.

ماجستير في العلوم

جامعة الملك فهد للبترول والمعادن

الظهران – المملكة العربية السعودية

ديسمبر ١٩٩٨

Chapter 1

Introduction

One of the challenging air defense problems is the ability of low flying helicopters to penetrate the air defense system, undetected by conventional radar. One way to detect this low-flying helicopter is to exploit its sound signature. Sound signature produced by helicopter noise may be employed for helicopter detection and recognition application. In this chapter: first a general picture of the recognition system is given. Then the characteristics of the feature encoder and classifier are described. In the following sections, both the motivation for and objectives of this thesis work are given and finally the outline of the work is presented.

1.1 Fundamentals of Recognition System

Recognition and classification of helicopters from its sound fall under the general problem of pattern classification. A pattern recognition system aims to recognize an object based on its previous knowledge of it. Such a system operates in three phases:

- a training phase, where, a set of parameters of the model is estimated so that in some sense the model "learns" the correspondence between the features and the labels of the objects. In all cases, the learning phase necessitates efficient learning algorithms for providing the system with truly representative reference patterns.
- a testing phase, in which the parameters of the model are then adjusted using a set of cross-validation data to achieve a good generalization of the system's performance. Cross-validation data consists of features and labels not included in the training.
- a recognition phase, during which an unknown feature with an unknown label is passed through the system which assigns a label at the output.

A pattern recognition system basically consists of a front-end feature extractor and a classifier as shown in Fig.(1.1).

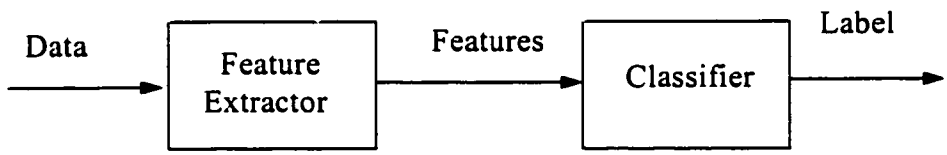


Fig.(1.1): Structure of a pattern recognition system

The feature extractor normalizes the collected data, removes irrelevant information, enhances interclass similarities and transforms them to the feature space[1]. In the feature space, the data is compressed and represented in such an effective way that objects from the same class behaves similarly and a clear distinction among objects from different classes exist. The classifier takes the features computed by the feature extractor and performs recognition.

1.2 Feature Extractor

One of the key assumptions made in the design of the helicopter recognition system is that, the signal can be regarded as wide sense stationary (i.e., the spectral characteristics are relatively constant) over a short interval of time, say 0.05 or 0.5 second (shown in Appendix A). The prime function of the front-end parameterization stage is to divide the input audio signal into blocks, which is digitized prior to analysis. From these blocks or frames, a smooth spectral estimate is obtained.

The dominant components in helicopter acoustic signal are the discrete spectrum generated from the blade passing frequency of the main rotor (3-40 Hz), the discrete spectrum from the rear rotor (25-110 Hz), the broad band noise generated from the laminar flow and vortex wake (200-600 Hz), noise emitted by engine drive train components (60-100 rotor angular speed), power turbine shaft N2 and the gas generator/compressor turbine N1 (900Hz - 2kHz), in addition to the effect of ground

reflections and the atmospheric attenuation. A typical time domain helicopter signal and its spectrum is shown in Fig.(1.2).

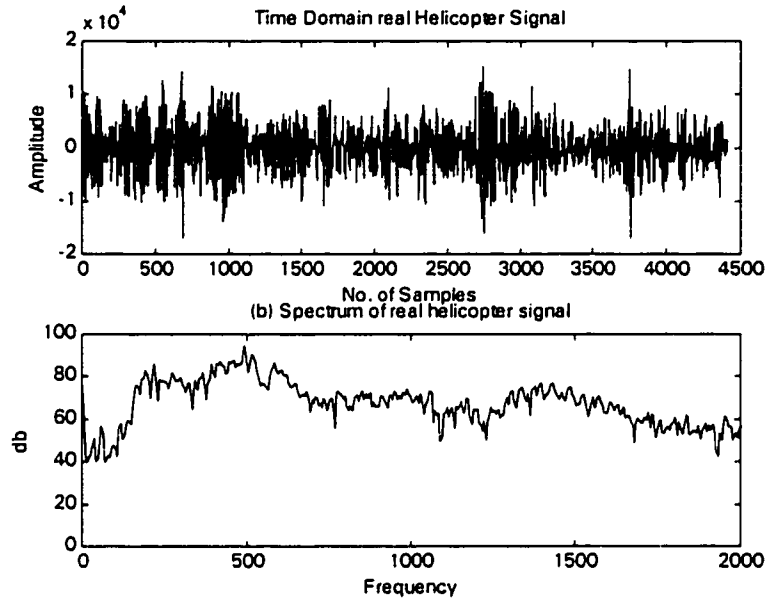


Fig.(1.2):(a)Time domain helicopter signal (b)Spectrum of helicopter signal

Therefore a low pass filter of cut-off frequency 2 kHz filters the signal. This pre-processing step removes some irrelevant information. This is followed by windowing, which involves multiplying the signal by a finite duration smooth window. The type of window chosen (shape and duration) influences the time and frequency resolution. The window used in this work is the Hamming window defined by the following equation and shown in Fig.(1.3).

$$h(n) = 0.54 - 0.46\cos\left(2\pi\frac{n}{N}\right) \qquad |n| \le N/2$$
$$= 0 \quad otherwise \qquad (1.1)$$

Where N is the window length.

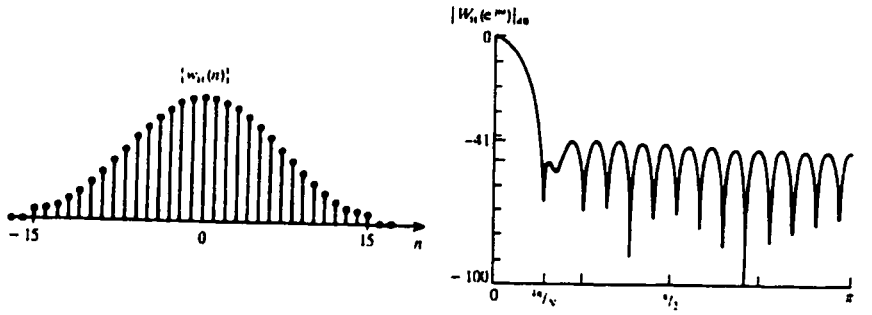


Fig.(1.3): a) Hamming window b) Fourier transform of Hamming window

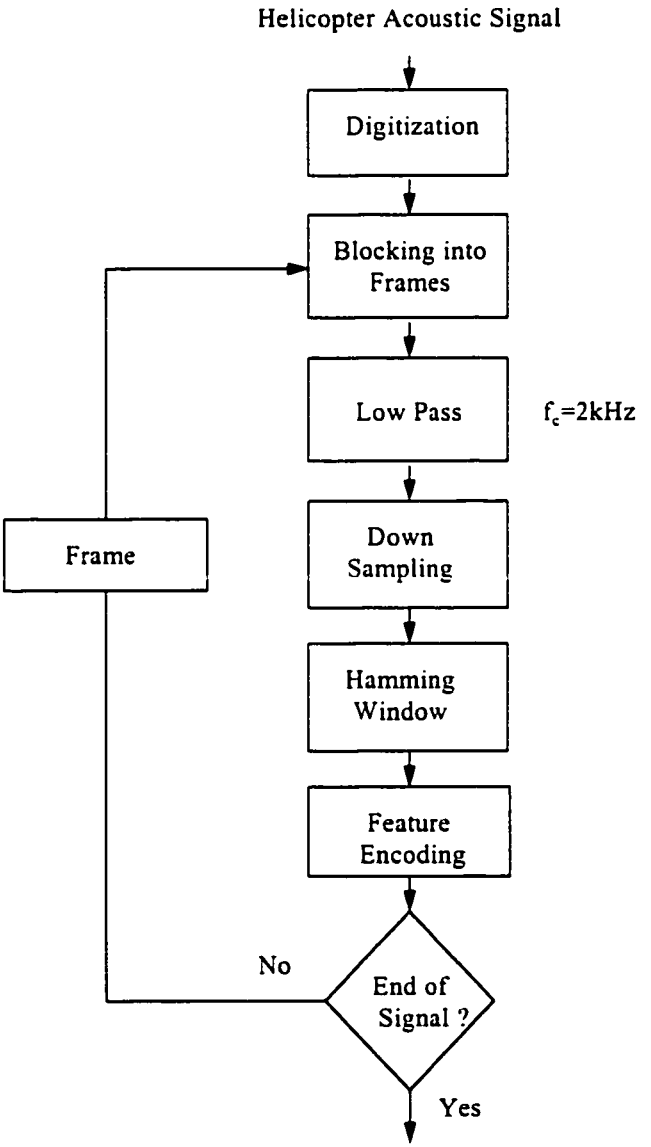


Fig.(1.4): Block diagram for feature extraction

Feature vectors are obtained from this windowed signal. The sequence of operations in converting the helicopter acoustic signal into set of parameters suitable for recognition process is shown in Fig.(1.4).

1.3 Feature Selection

An important step in the signature identification and recognition process is to extract adequate information for good discrimination, and, at the same time to capture the information in a form and size that is amenable to effective modeling. The helicopter noise is usually digitized at a rate of 4 kHz or higher, using 8 bits or more per sample, requiring tens of thousands of bytes for a few seconds of the signal. Whereas these large amounts of information are needed to characterize the helicopter sound, the essential characteristics of helicopter sound changes relatively slowly, permitting a representation requiring significantly less data. Helicopter noise can be parameterized over relatively long time periods of 0.05 to 0.5 second called frames. Considering a 0.1 second frame length and a 10 dimensional feature vector per frame (say), this corresponds to a data reduction ratio of 400/10=40 at 4kHz sampling rate. The process of reducing data while retaining classification information falls under the general heading of feature extraction.

To get the spectral estimates, in the form of feature vectors, linear predictive coding (LPC), Fourier analysis and a number of other additional transformations, such as, reflection coefficients, LPC cepstrum coefficients, line spectral pairs (LSP)

can all be employed. Wavelet transforms, the recently proposed technique for signal processing, can also be applied for signal de-noising prior to feature extraction.

1.4 Classifiers

The most common methods for classifiers used in acoustic signal recognition are:

- Dynamic programming
- Stochastic modeling
- Artificial neural networks

Although these methods have their own intrinsic characteristics, they nevertheless share some common ones. The techniques are, to a large extent, independent of the type of acoustic signal. The accuracy of such systems is highly dependent on the amount of training data, in general, the more training data the higher the accuracy.

Dynamic programming is based on optimality principle. To take into account physical aspects of the problem and to limit the number of warping paths considered, it is necessary to impose some constraints on the warping functions like endpoint detection, monotonicity, local continuity and slope constraints [2]. Even with these constraints, dynamic programming is very time consuming.

Stochastic modeling is explicitly based on the statistical characteristics of the signal. These approaches belong, to a large extent, to the statistical decision theory. It is based on minimizing the average loss per decision as small as possible. Hidden

Markov models (HMM), Discrete HMM and Continuous Density HMM are a few examples of stochastic modeling. In HMM, "exponentially decaying" state duration density is inappropriate for acoustic signal representation. Therefore, it requires explicit modeling of duration density which is computationally very expensive.

In recent years, Artificial Neural Networks (ANNs) have been widely applied for pattern classification (e.g., automatic speech recognition). The capability of learning from examples, producing arbitrary non-linear functions of input, and the highly parallel and regular structure of ANNs make them especially suitable for hard classification. The feature vectors are applied as input to the network. At the training stage, the network adjusts its variable parameters (synaptic weights) to capture the features of the object. The back propagation algorithm can be used for network training of the network.

1.5 Motivation and Objectives

Helicopter can act as an intruder's transport, a highly mobile weapon platform and as an escape vehicle after an intrusion has been committed. This is because of their low-flying ability to penetrate the defense system, undetected by conventional radar. Building a system of remote sensors to detect and track single and multiple very low-flying helicopters is an important defense problem. Detection of helicopters using their sound signatures has been a focus of a number of research works[3-11]. Conventional methods use the ratio of the main and tail rotor frequencies and their

harmonics to extract the helicopter noise features. Moreover they have used simulated spectrum based only on a fixed discrete spectrum from the rotor, without considering the effect of the directivity on the variability in the harmonic distribution of the main rotor noise.

A simulator designed for this work includes more realistic helicopter noise characteristics, including aerodynamic vortex shedding, blade thickness, atmospheric attenuation and terrain effects. It is interesting to investigate speech-coding algorithms to get the spectral parameters of the signal. In this thesis work, spectral parameters are calculated through linear predictive coding techniques. Different spectral parameters, such as LPC coefficients, reflection coefficients, LPC cepstrum, line spectral pairs (LSP) are used. In one specific case, wavelets are also used for de-noising and then spectral estimates are obtained.

In this work an ANN is used as a pattern classifier. Most other classification methods are based on statistical information, where a number of assumptions are made for mathematical interpretation of the physical phenomenon, like linearity, stationarity and Gaussianity. Any deviation from these assumptions, in real life, results in suboptimal performance. ANN has the ability to learn from examples, eliminating the need for a mathematical model. Due to the use of non-linear functions, they can model the underlying non-linearities in the physical process. So it is expected that the ANN classifiers, along with acoustic signal encoding techniques, will preserve the essential sound characteristics which in turn will produce better detection and classification results.

For each feature extraction technique, the performance of recognition system will be compared in terms of the classification accuracy. In each of them the effect of network size (number of neurons), number of training patterns and feature vector length on the classification accuracy will be studied.

1.6 Outline of the Thesis

Chapter 1 gives a brief introduction of the recognition system along with the motivations and objectives of the thesis work. Chapter 2 presents a relevant literature review. Chapter 3 is dedicated to the helicopter sound analysis and simulation. The aerodynamics effects, broad band noise, Doppler's effect, atmospheric attenuation and ground reflection issues are presented in this chapter.

Extraction of feature vectors using different encoding techniques is presented in chapter 4. LPC coefficients, reflection coefficients, LSP's and LPC cepstrum coefficients with their characteristic features are all included. Application of Wavelet transform for noise reduction is the topic of chapter 5.

Chapter 6 introduces neural networks and their use in signal processing and for a classifier design. Results of helicopter detection and classification are presented and analyzed in chapter 7 and 8 respectively. Chapter 7 and 8 further include comparison of different encoding techniques. Chapter 9 includes conclusions and recommendations.

Chapter 2

Literature Review

Summary

In this chapter a review of the different techniques for helicopter detection and classification is given. Helicopter acoustic signal has a distinctive impulsive nature due to blade rotation, which could be exploited to design a recognition system. Conventional classifiers for helicopter identification were based on statistical information of the main and tail rotor frequencies. An ANN has been used as a classifier in this detection problem due to their speed and ability to work in the environment involving incomplete or missing information. With the use of an ANN, more information about the spectral features is used in training the classifier. Some researchers have compared the use of ANN in the classifier design to the conventional classifier in helicopter detection problem.

The research efforts in the helicopter recognition area can be divided into two main categories:

- Use of main and tail rotor periodicities
- Use of an ANN as a classifier

2.1 Use of Main and Tail Rotor Periodicities and Harmonics

A conventional method of helicopter identification uses the ratio of main and tail rotor frequencies of the helicopter acoustic signal and compares it with a table of known value [10]. This method works as long as the tail rotor noise is present at the receiver. However if the helicopter does not have a tail rotor, or the high frequency tail rotor acoustic signal is attenuated due to orientation of helicopter and/or atmospheric conditions, then this method performs poorly.

Mori et. al. [11] investigated the use of a network of remote sensors to detect and track, single and multiple very low-flying helicopters where information from different sensor nodes was exchanged to improve the resolution of the sound detectors. The overall recognition performance is measured in terms of probability of false alarm and probability of missing. The fusion of data from multiple sensors was necessary to overcome the relatively low performance of the detector.

Cecil [5] has discussed the requirement and characteristics of an optical microphone that would scan the environment in the vicinity of the helicopter and would produce a signal associated with specific acoustic disturbances. Contrary to conventional microphones, optical microphones have superior directionality resolution and, if properly insulated, will not be confused by other acoustic signals. Experiments have demonstrated that an acoustic radiator impresses upon the atmosphere a fundamental frequency and harmonics. This phenomenon frequency shifts the Doppler frequency resulting from convective air motion. The net effect is the creation of a signature that depends upon the acoustic pressure. A systematic method, namely, the location of maximum acoustic pressure, can be implemented for helicopter location, as the pressure is higher near the source. Cecil has conducted his experiment in laboratory with a small speaker as the acoustic radiator and a reference beam laser Doppler system to measure the frequency associated with particle vibration and convection. The signal was Fourier transformed and the results were compared to an analytical model. Acoustic frequency between 100Hz and 400Hz were measured with a typical accuracy of 5%.

Feder [6] has applied parameter estimation approach for helicopter signal extraction observed with a wideband interference. The helicopter signal can be characterized as periodic or almost periodic. The basic problem is the estimation of the period or the fundamental frequency. He has used a model-based approach. The periodic signal is defined up to some parameters; one of them is the period. The other parameters describe the deviation from strict periodicity. The wideband signal is

modeled as a Gaussian random process with an unknown spectrum. To simplify the analysis an autoregressive model was assumed, whose parameters determine the bandwidth and the center frequency of the signal. The likelihood of the observed signal can be maximized to produce the estimates of parameters. He used an iterative algorithm (expectation-maximization algorithm) for maximizing the resulting likelihood function.

Vezzosi and Ehrmann [8] used the idea of applying far field acoustic signal produced due to the main rotor. They proposed a mathematical model that leads to a sufficient description of the low-frequency part of the signal. The acoustic signal was assumed to be produced by a point-like rotating source with periodic excitations. They have measured the harmonics and defined a receiver, which uses both the modulus and phase of the measured harmonics. The receiver allows helicopter classification according to the number of blades and Mach number. The receiver may be considered as the variant of the matched filter to improve its detection and classification abilities.

The rotation of the helicopter blades produces stochastic signal for which appropriate probabilistic models exhibit periodically time-variant parameters and hence it is called cyclostationary process. Due to the presence of harmonic spectral lines in the power spectral density (PSD), helicopter signals are closely related to first order periodicity. Zhiping Lin [7] has applied cyclostationarity for helicopter signal detection without imposing any restriction on the additive noise. He has also used the cyclic frequency smoothing method, useful for acoustic signal processing, to

have a better estimate of the average fundamental frequency of time-variant Doppler shifted helicopter acoustic signal.

2.2 Use of ANN as Classifier

A helicopter's acoustic signal is a product of a complex phenomenon and any correlation between this signal and the helicopter type is also complex. An ANN, which can learn any complex function [12], can be used as a solution. ANNs are used due to their processing speed. They are a collection of simple computational elements. Large assemblies of these simple elements can solve problems requiring massive constraint satisfaction. They learn from input-output data, eliminating the tedious programming associated to complex problems.

Cabell et. al. [4] have trained a neural network classifier for helicopter detection and compared its performance with a conventional recognition system. The conventional method relies on the ratio of main and tail rotor blade passage frequencies. They have designed a custom main/tail rotor ratio identification system called HARMO system. This system has used frequency domain information to distinguish between helicopters. Harmonic detector is used to locate the fundamental frequency of the main rotor. From the main and tail rotor frequency ratio, the tail rotor frequency is determined. The highest tail rotor harmogram value identifies the corresponding helicopter. For neural network classifier, in addition to above two features, first and third order curves were fitted to the peaks of first eight harmonics

of the main rotor noise. The slope of the best fit-line and coefficients of x, x^2 *and* x^3 terms in best-fit cubic curve were used as features. They have used two prototype spectra. The conventional method has outperformed the ANN classifier when main and tail rotor noises were present, but the conventional method could not identify helicopters when tail rotor noise was removed from the spectrum. At 20dB signal-to-noise ratio (SNR), ANN classifier has correctly identified the helicopter 100% of the time as compared to 50% employing conventional method, in the absence of tail rotor noise.

In a later paper, Cabell et. al. [9] have presented the results of a preliminary study of a new method of helicopter identification. They have used the characteristics of the main rotor noise to improve the performance of the identification system. They have developed two pattern classifiers: a statistically based Bayes classifier and an ANN classifier. The Bayes classifier used the mean vector and the covariance matrix as features to identify unknown patterns. They have selected peaks of spectral amplitudes of fundamental and first seven harmonics and fitted first and third order curves using least square regression (in the same way as they have done in their previously described work [4]). The slope of the fitted-line and the coefficients of the first and third order terms from the fitted cubic polynomial were used as features for ANN classifier. They employed seventeen recordings of three different helicopter types at different location from NASA database. The Bayes classifier used, partitioned the feature space according to distribution of feature values. The classifier was used under the assumption that the class populations are normally distributed

which was supported by the performance. The Bayes classifier has identified 67% of noise segments of three helicopters while ANN classifier was correct 65% of the time.

El-Shafei and Ahmed [3] have used a recorded helicopter sound and a number of non-helicopter sounds to train their ANN classifier. They have noticed that a strong peak at engine vibration frequency characterizes the spectrum of a helicopter sound, side lobes at 10% to 25 % of the main peak and components in the frequency band 150Hz to 350Hz. Goertzel algorithm (GFFT) was used for evaluating the spectrum at selective frequencies in the band of 160-350 Hz. Discrete Fourier transform (DFT) feature vectors of length 30 were used as input vector for the training and testing of classifier. ANN classifier with one hidden layer of neurons was used. The results obtained have shown 99.5% correct detection of helicopter over test samples.

Chapter 3

Helicopter Sound Analysis

Summary

Helicopter sound analysis deserves special consideration. Apart from propeller noise (due to the main rotor), non-axial inflow conditions of tail rotor, physical dimensions of blades, position and velocity vectors of observer and aircraft, atmospheric attenuation and distortion and ground effect make significant contributions to the overall received sound signal. The purpose of this chapter is to examine the physical characteristics and theoretical prediction of helicopter noise. These notions are embodied in a computer program to yield a helicopter sound with realistic details. This simulator has been used to generate different flight conditions for training and testing of neural network-based techniques for detection and classification of helicopter sound.

Sources of Helicopter Sound

For helicopter identification and classification problem, it is almost impossible to train the system using real sound samples under all possible flight conditions. Due to unavailability of these direct measurements, the only method of accessing the impact is to utilize a reliable simulator, which makes use of directly relevant data e.g., rotor angular velocities, number of blades etc. or on some empirical prediction procedure. To be successful, acoustic generation must be based on reliable definition of helicopter performance. Schematic of a single rotor Helicopter is shown in Fig.(3.1).

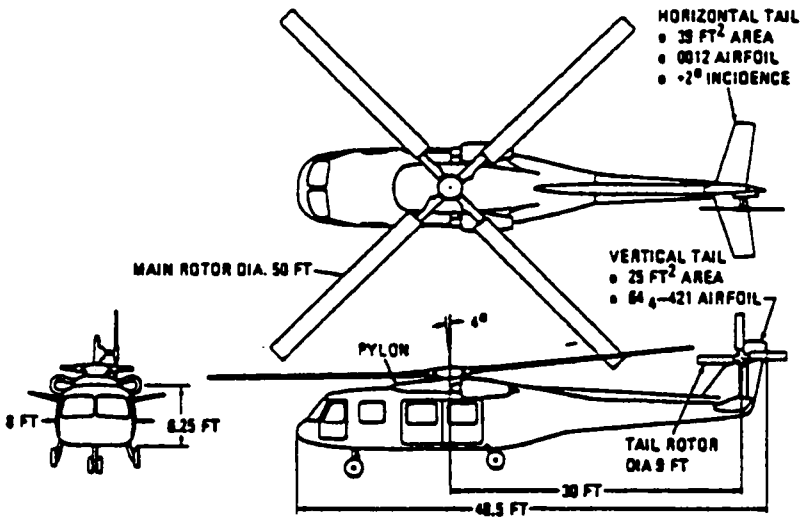


Fig.(3.1) Schematic of a Helicopter

The simulator used in this thesis work consists of following four parts:

1. Helicopter database, which contains the key physical parameters of each helicopter of interest.

2. Sound production model.
3. Flight trip simulator, which includes such variable parameters as speed, initial observation point, and 3D-flight path with respect to the observer.
4. Disturbance model, which includes fading, Doppler shift due to wind gusts, and sound interferences. e.g., jet engine noise.

The quantitative prediction of helicopter sound is extremely difficult. Few sources of lifting-rotor (and tail rotor) are identical to those associated with the propeller. A significant additional phenomenon known as 'blade slap or thump' [13,14] is also an additional noise source. It occurs when one of the rotors passes through wake or tip vortex created by another blade. Tail rotor noise, which varies from mid to high frequencies, is added to main rotor noise. No two-helicopter types produce the same sound field [15]. Helicopter rotor noise tends to grow in intensity and character, as tip speed and power increases, but the rotor noise effects are significantly asymmetric [16]. For a non-hovering helicopter, there will always be a substantial difference in relative tip speed between the blade on the "advancing" side and that on the "retreating" side of rotor disc plane with respect to the flight direction as shown in Fig.(3.2).

High speed of the advancing blade produce the asymmetric noise output often preferentially directed ahead of the helicopter [15]. The same effect occurs, albeit to a lesser degree, with the tail rotor.

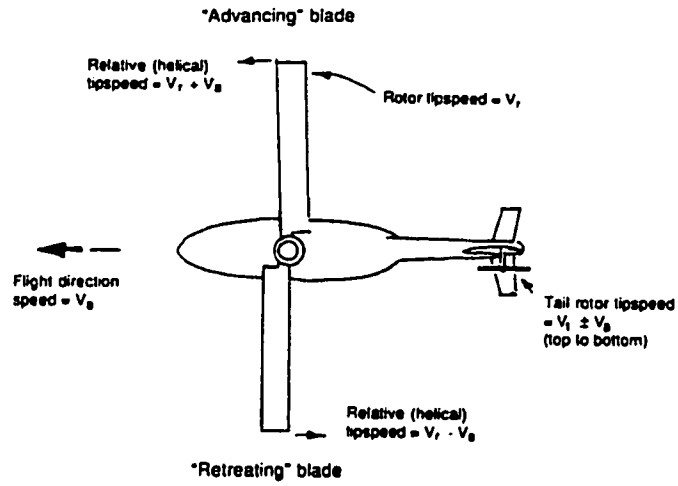


Fig.(3.2) Asymmetry in the Aerodynamics of Helicopter Blade

At low blade speeds, the mean steady loading can be used to predict the fundamental rotor tone and low-order harmonics, using the propeller analogy. But at higher helical tip Mach numbers, thickness noise increases rapidly and propeller analogy breaks. Atmospheric attenuation and ground reflection effects may further complicate measurements. The observed sound is a function of the angular velocities of main and tail rotors and their number of blades, physical dimensions of blades, the aircraft position vector with respect to observer, the velocity vectors of the aircraft and the observer, atmospheric attenuation and disturbances, and the ground terrain.

A typical section of the helicopter acoustic signal is shown in Fig.(3.3). It is largely a tonal signature, a combination of tones from main rotor (3-40 Hz), the relatively higher tones due to tail rotor (25-110 Hz), broad band noise generated from the laminar flow and vortex wakes (200-600 Hz), noise emitted by engine drive train components (60-100 rotor angular speed), power turbine shaft (N2), and the gas

generator/compressor turbine (N1) (900-2000 Hz), in addition to the effect of ground reflection and the atmospheric attenuation.

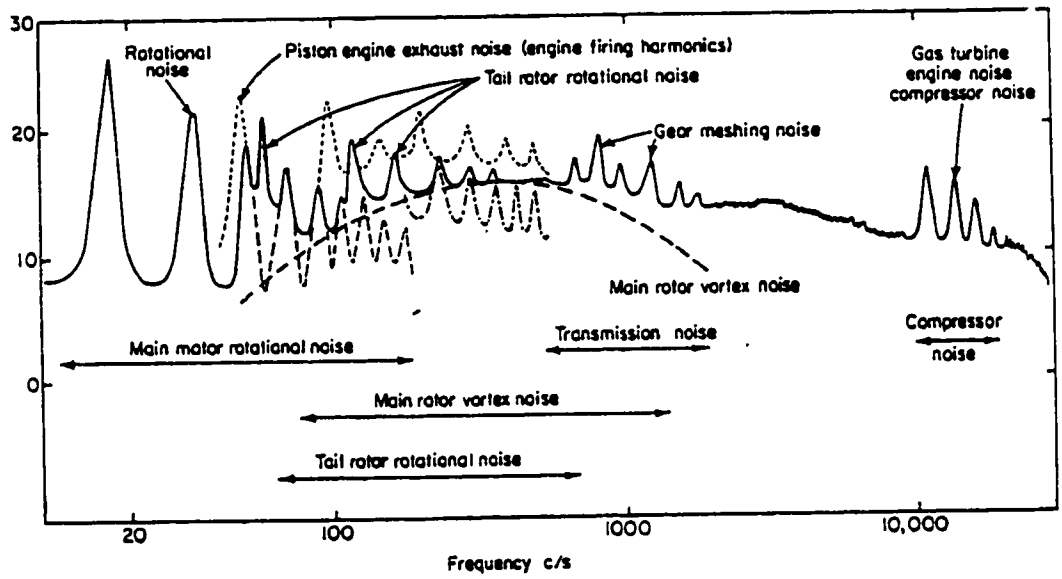


Fig.(3.3) Helicopter Noise – narrow band spectrum

The large number of separate sources contributing to helicopter noise can be identified and summarized as:

1. Rotor noise, due to main and tail rotor, this consists of (a) rotational noise (b) vortex or broad band noise (self-noise) (c) discrete frequency noise known as 'blade slap'.
2. Noise emitted from engine drive train components, power turbine, shaft (N2) and gas generator/compressor noise.
3. Effect of ground reflection and atmospheric attenuation

3.1 Rotor Noise

3.1.1 Main Rotor Noise

The noise from rotating blades is of dipole order, and is generated by the fluctuating pressure distribution on the surface of the blade. Lighthill [17] has shown that a fluctuating point source behaves like a dipole, giving source strength proportional to the magnitude of the force resolved in the direction of the observer. Force fluctuations relative to the observer are the result of unsteady aerodynamics effects and blade rotation. So rotor noise increases rapidly with the tip speed.

In Fig.(3.3), the presence of a large number of peaks suggests that a strong source of propeller noise arise from periodic excitation at the blades. Peaks are present at the fundamental and successive harmonics of the blade passage frequency. In addition to the discrete frequency component of the spectra, broad band noise is found in a region of 1-2 kHz, which, at low tip speeds, becomes comparable in magnitude to the discrete frequency noise. This broadband noise is due to random disturbances at the blades.

3.1.1.1 Rotational Noise

Analysis of the rotational noise was first presented by Gutin [18]. He considered the noise generated by the steady load on a propeller. Lowson and Ollerhead [19] and Barry and Moore [20] employed periodically varying loads in their studies of rotating sources.

In most of the propeller noise prediction schemes in use, the noise source on the blades is modeled by a single point source at 80% of the blade span [21]. Its justifications are velocity dependence, which causes a strong peak in source strength close to tip, and reduction in complexity of the computations required for a complete span-wise integration [22].

The most important rotational noise sources are the steady and fluctuating blade loads. The sound pressure, $p(x_i,t)$, generated by a fluctuating dipole point source is [23]

$$p(x_i,t) = -\frac{\partial}{\partial x_i}\frac{F_i(t-r/c_0)}{4\pi r} \tag{3.1}$$

where $x_i = (x_1,x_2,x_3)$ represents the observer location, $F_i(t)$ the fluctuating force component in each direction, c_0 is speed of sound and r is the distance of the assumed point source from the point of observation.

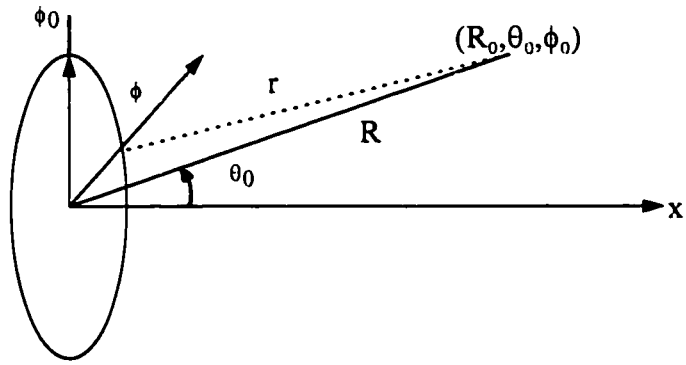


Fig.(3.4):Coordinates of rotor blade and observer

For rotor blades, the loads may be resolved into their lift *L(.)* and drag *D(.)* components, which are assumed to act at a point displaced by a distance *a* from the rotor axis. The drag force is in the plane of the rotor disc and the lift force is perpendicular to the plane. The acoustic field generated by these two forces can be expressed in spherical coordinates as [23]

$$p(r,\theta_0,\phi_0,t)=\frac{\partial}{\partial x}\frac{L(t-r/c_0)}{4\pi r}-\frac{1}{a}\frac{\partial}{\partial \phi_0}\frac{D(t-r/c_0)}{4\pi r} \tag{3.2}$$

where θ_0 is the angle between the axis of rotation and the radius vector *R*.

To explain the effect of blade rotation, the rotor is modeled by a continuous distribution of sources on a ring of radius *a*. The strength on this ring is zero every where apart from at the instantaneous locations of each blade, where it is equivalent to the instantaneous blade loading. As the rotor has an angular velocity Ω, the source strength on the stationary ring will vary with time. At a fix point ϕ, the source strength variation will consist of a series of impulses which occur every time a blade passes that point. The contribution of the acoustic field from the source at ϕ from a rotor with B blades will be [23]

$$dp=\frac{\partial}{\partial x}\frac{L(\phi)}{4\pi r}\sum_{m=-\infty}^{\infty}\delta\left(\phi+\frac{2\pi m}{B}-\Omega(t-r/c_0)\right)-\frac{1}{a}\frac{\partial}{\partial x}\frac{D(\phi)}{4\pi\phi_0}\sum_{m=-\infty}^{\infty}\delta\left(\phi+\frac{2\pi m}{B}-\Omega(t-r/c_0)\right) \tag{3.3}$$

where $\delta(\phi)$ is a Dirac delta function [23].

Applying the Fourier series expansion to above equation will result

$$dp = \frac{B}{8\pi^2}\sum_{m=-\infty}^{\infty} e^{-imB\Omega t}\left[L(\phi)\frac{\partial}{\partial x} - \frac{D(\phi)}{a}\frac{\partial}{\partial \phi_0}\right]\frac{e^{imB(\phi+\Omega r/c_0}}{r} \tag{3.4}$$

To get the acoustic field generated by complete ring source, the above equation is integrated over ϕ, and the variation of lift and drag as a function of position is accounted for. Due to inflow distortion of some type, loading on the rotor varies periodically with position and the effect of this can be incorporated by expanding $L(\phi)$ and $D(\phi)$ in a series of rotor loading harmonics. It is assumed that both lift and drag vary in the same way to a given inflow distortion and may be expanded as:

$$L(\phi) = L_0\sum_{n=-\infty}^{\infty} A_n e^{in\phi}, \quad D(\phi) = D_0\sum_{n=-\infty}^{\infty} A_n e^{in\phi}$$

where A_n is the amplitude of n^{th} rotor loading harmonic.

So acoustic field from the complete ring source is:

$$p(r,\theta_0,\phi_0,t) = \frac{B}{8\pi^2}\sum_{m=-\infty}^{\infty} e^{-imB\Omega t}\sum_{n=-\infty}^{\infty} A_n\left[L(\phi)\frac{\partial}{\partial x} - \frac{D(\phi)}{a}\frac{\partial}{\partial \phi_0}\right]\int_{-\pi}^{\pi}\frac{e^{i(mB+n)\phi+imB(\Omega r/c_0)}}{r}d\phi \tag{3.5}$$

To evaluate the acoustic far field, the propagation distance r may be approximated by

$$r \cong R - a\cos(\phi - \phi_0)\sin\theta_0$$

So in magnitude and phase terms, 1/r is replaced by 1/R.

Defining integral over ϕ in terms of Bessel function, replacing $\frac{\partial}{\partial x} = \cos\theta_0\frac{\partial}{\partial R}$

and incorporating the source Mach number $M_s = \frac{\Omega a}{c_0}$ the acoustic field becomes:

$$p(R,\theta_0,\phi_0)=\frac{B}{8\pi Ra}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}A_n\, e^{-imB\Omega(t-R/c_0)+i(mB+n)\phi_0}\times$$
$$\left[mB\,M_s\,L_0\cos\theta_0-(mB+n)D_0\right]i^{mB+n+1}\,J_{mB+n}(mBM_s\sin\theta_0) \tag{3.6}$$

Bessel function, $J(\xi)$, tend to zero when their argument $mBM_s\sin\theta_0$ are very much less than their order (mB+n). So, far field radiation is dominated by the terms in the series for which

$$M_s\sin\theta_0>\left|\frac{mB+n}{mB}\right|$$

In Eq.(3.5) each term is rotating with a phase speed $\Omega_p=\frac{mB\Omega}{mB+n}$ and so efficient radiation occurs only when

$$M_p\sin\theta_0=M_s\left|\frac{mB}{mB+n}\right|\sin\theta_0>1$$

where M_p is the pattern Mach number $\Omega_p a/c_0$.

The tip velocity variations among various aircraft are limited to Mach number between 0.6 and 0.66 due to aerodynamics reasons. For a subsonic rotor above condition will not be satisfied unless -*n* is of the same order of magnitude as *mB*.

We can consider two cases:

- first, when the rotor is highly loaded and has a clean inflow, only steady loading terms (n=0) need be considered, and the amplitude of each blade passing harmonic in the noise spectrum will be:

$$p_{mB}(R,\theta_0,\phi_0)=\frac{mB^2\Omega}{4\pi Rc_0}\left(L_0\cos\theta_0-\frac{D_0}{M_s}\right)J_{mB}(mBM_s\sin\theta_0) \tag{3.7}$$

- alternatively if the blade encounters a discrete velocity disturbance which increase the lift and drag on the blades to the extent that it dominates the force fluctuations over the rotor plane, the signal generated will be very impulsive. The expression for the sound pressure due to m^{th} harmonic is given by

$$p_{mB}(R,\theta_0,\phi_0) = \frac{mB^2\Omega}{4\pi R c_0}\sum_{n=-\infty}^{\infty} e^{in\phi}\left(L_0\cos\theta_0 - \frac{D_0(mB+n)}{M_s mB}\right) J_{mB}(mBM_s\sin\theta_0) \qquad (3.8)$$

Lawson and Ollerhead [19] presented a prediction scheme using above equation in which they assume $D_0 / L_0 = 0.1$ and $BL_0 = L_T$ = total thrust generated by the rotor. The estimated mean square of the each harmonic of the lift is found to decay at a rate=$1/n^{2.5}$. The pressure at the observer can then be approximated by:

$$p_{mB}(R,\theta_0,\phi_0) = \frac{mBL_T\Omega}{4\pi R c_0}\left[\sum_{n=-\infty}^{\infty}\left(\cos\theta_0 - \frac{0.1*(mB+n)}{M_s mB}\right)^2 \frac{J_{mB}^2(mBM_s\sin\theta_0)}{n^{2.5}}\right]^{1/2} \qquad (3.9)$$

The time varying harmonic component is then

$$p_{mB+n}(R,\theta_0,\phi_0,t) = p_{mB+n}(R,\theta_0,\phi_0)e^{-i(mB+n)\Omega(t-R_0/c_0)+i(mB+n)\phi_0} \qquad (3.10)$$

In our simulation we are assuming low tip speed (Mach number 0.6-0.66), so we can approximate Bessel function by its asymptotic value given as [23]

$$J_{mB}(mBM_s\sin\theta_0) \cong \frac{(mBM_s\sin\theta_0)^{mB}}{2^{mB}\,mB!}$$

$$\cong \frac{e^{mB}}{\sqrt{2\pi mB}}\left[\frac{M_s\sin\theta_0}{2}\right]^{mB} \qquad (3.11)$$

Let us denote

$$\tilde{P}_{mB} \cong \frac{e^{mB}}{\sqrt{2\pi mB}}\left[\frac{M_s \sin\theta_0}{2}\right]^{mB}$$

so pressure variation at the observer is

$$P_{mB}(R,\theta_0,\phi_0) \cong \frac{B\Omega}{4\pi R c_0}\left(L_0 \cos\theta_0 - \frac{D_0}{M_s}\right)\tilde{P}_{mB} \tag{3.12}$$

3.1.1.2 Broad Band Noise (Vortex Noise)

The broadband noise is due to two significant sources. One, from the shedding of vorticity at the airfoil trailing edge when the blade is operating in smooth airflow, induced due to local surface pressure fluctuations on the blade. The other, when the blades move in a turbulent airflow. The intensity of the broad band noise is given by [24,25]

$$I = k(R_e)^{-0.4} A \rho_0 \frac{V_T}{c^3} \cdot \frac{\cos^2\theta_0}{r_0^2} \tag{3.13}$$

where k is of the order 10^{-4}, and Reynolds number R_e is based on tip speed V_T and mean chord length c, A is the blade area.

The constant k in the above expression is not easily calculated. Davidson and Hargest [26] have used actual helicopter measurements. Their relationship for the sound pressure level of the vortex noise in terms of the blade parameters is:

$$Max.S.P.L. = 60\log_{10} V_T + 20\log_{10} C_{Lt} + 10\log_{10} S - 84.0 \quad dB \tag{3.14}$$

where V_T = blade tip speed (ft/s)

C_{Lt} = lift coefficient referred to blade tip

S = total blade span area of the rotor (ft^2)

Goddard and Stuckey [27] obtained a more general relationship, which gave the overall S.P.L. as

$$O.A.S.P.L. = 16.6 \log T + 26.8 \log V_T - 20 \log r - 20 \log \sec \phi + 2.8 \quad dB \qquad (3.15)$$

where T = thrust per blade (lbs)

r = distance from source to observer (ft)

ϕ = angle between tip located dipole axis and observer

If the peak frequency f_p is defined in terms of Strouhal number, S_t, vortex spectrum shape is very well defined

$$f_p = S_t \frac{V_T}{c_0} \qquad (3.16)$$

The broadband noise increases as V_T^3 for low, V_T^5 for medium and V_T^8 for high tip speeds [21].

In our simulation, the lower cutoff frequency is taken as $f_L = {f_p}/{Q}$ and higher cutoff frequency is $f_H = f_p Q$ with $f_{max} = f_p$ and Q is 1/8 called quality or shape factor.

3.1.1.3 Blade Slap

Blade slap or blade bang is the sharp cracking or banging sound associated with the main rotor under certain operating conditions. It is impulsive in nature and is the most objectionable and annoying. It is agreed that this noise is produced when one

blade passes through the tip vortex shed by the previous or some other blade, but actual mechanism of generation is not fully understood [23].

Bell Helicopter Co. [28] suggests that at low speeds it is caused by the rapid change in relative angle of incidence of the blade as it encounters the wake of a previous blade while at high speed it is more likely to be due to local shock waves on the advancing blades. In the work at Southampton, blade slap was considered to arise from blade/vortex interaction. It was found that the noise was proportional to V_T^6 and square of the vortex strength.

3.1.2 The Effect of Physical Parameters

3.1.2.1 Number of Blades

The effect of increasing the number of blades is to increase the fundamental frequency and to cancel all other harmonics except those which are integral multiples of the blade number. An increase of blade number tends to increase the intensity of vortex noise. The unsteady force fluctuations on each blade are random in nature, so that the total noise output will be proportional to the number of blades.

3.1.2.2 Tip Speed

Hubbard and Lassiter [29] found that, the measured overall noise (which include all components) increases rapidly with increasing tip speed in the subsonic range. With the increase in tip speed the harmonics of the blade passage frequency increases at a

greater rate than fundamental. Another effect of the tip speed is to change the directivity pattern.

3.1.2.3 Forward Speed

The effect of the forward speed is to modify the length of the acoustic path between the source point and the observer. Expression(3.10) is valid for helicopters in hover state. For moving helicopter the distance can be substituted by $R_0 = t * v$, where v is the component of the helicopter velocity in the direction of the observer. So Eq.(3.10) can be modified as:

$$p_{mB+n}(R,\theta_0,\phi_0,t) = p_{mB+n}(R,\theta_0,\phi_0)e^{-i(mB+n)\Omega(1-v/c_0)t+i(mB+n)\phi_0} \tag{3.17}$$

The above equation shows that the movement of helicopter induces a Doppler effect in the form of frequency shift in the observed spectrum. This change of m^{th} harmonic is given by $\Delta f_m = \frac{mB\Omega v}{2\pi c_0}$. The change in the orientation of the helicopter with respect to the observer causes a continuous fluctuation in the relative magnitude of each of the component of the received signal. These two effects result in jitters in the observed spectrum.

3.1.3 Tail Rotor Noise

The effect of the rear rotor is usually much less than that of the main rotor. Moreover, it is highly sensitive to the aircraft orientation with respect to the observer. It consists of a sideways-facing rotor usually running at a constant speed. The effect of the rear rotor is maximum when its axis is in the direction of its wake. Because of

its smaller diameter, it runs at a much higher angular speed than the main rotor. As illustrated in Fig.(3.3), it produces a series of harmonics in a similar manner to the main rotor, but it has a higher fundamental frequency (40-120 Hz). On most helicopters it is the most noticeable and disturbing noise source, especially from high speed rotors. The vortex noise from tail rotor is generally of lower level than the discrete or rotational tones [21].

3.2 Noise Emitted from Engine

3.2.1 Transmission Noise

Gear and transmission noise is important internally; although for small helicopters this can be heard externally. The range of noise is usually between 500-1500 Hz, the most serious audible interference frequencies.

3.2.2 Power Plant Noise

3.2.2.1 Piston Engine

On large helicopters the engine is well isolated from the cabin and the noise is only an external problem. It occurs in the range of 200-1000Hz. Primary source of noise is the exhaust. The noise emanating from the periodic expulsion of the hot gases of combustion produces harmonics of the engine firing frequency.

3.2.2.2 Gas Turbine

The main source of noise on the turbine engine is the 'compressor whine' from the inlet. Additional gears are required to give a higher gear ratio in this type of engines. So transmission noise increases in turbine engines. The exhaust noise of a turbine engine is small because the exit velocity of the gases is relatively low and continuous. In case of fixed wing turbine aircraft, high frequency noises from the compressor are extremely annoying.

Due to lack of real data, the two vibration frequencies of the engine were fixed during simulation. The engine noise received, is modulated by the blade passing pressure and Doppler shifted due to motion of the helicopter.

3.3 Effect of Ground Reflection and Atmospheric Attenuation

Sound waves travel from source to receiver outdoors through an atmosphere that is in constant motion. Turbulence, temperature, wind gradients, viscous and molecular absorption and reflection from the earth's surface all affect the amplitude and create fluctuations in the sound received. The received sound power level $L_{p\theta}$, from a directive sound source of power L_w, is given by the expression [30]:

$$L_{p\theta} = L_w + DI_\theta - 20\log r - A_e - 11 \quad dB \tag{3.18}$$

where $L_{p\theta}$ = Sound pressure level at a receiver located in the direction θ, a distance r (m) from the source

DI_θ = Directivity index of source in the direction θ

r = Distance of receiver from the source (m)

A_e = Excess attenuation caused by the environmental conditions (dB)

Excess Attenuation

The excess attenuation owing to environmental and other conditions is the attenuation beyond that caused by wave divergence. It includes one or more of the following factors:

A_{e1} = effect of the difference in value of ρc from 400 mks rayls

$$\left(\rho c = \frac{p_{ref}^2}{I_{ref}} \quad having\ units \frac{N.s}{m^3} = mks\ rayls\right)$$

A_{e2} = attenuation by absorption in the air (dB)

A_{e3} = attenuation by rain, sleet, snow, or fog (dB)

A_{e4} = attenuation by barriers (dB)

A_{e5} = attenuation by grass, shrubbery and trees (dB)

A_{e6} = attenuation and fluctuations owing to wind and temperature gradients, to atmospheric turbulence, and to the characteristics of the ground (dB)

In the case of air to ground propagation, the presence of barriers, trees and plantings on the ground and the effect of ground created wind and temperature gradients are usually not important except for very low flight.

3.3.1 Effect Of The Difference In Value Of ρc

For the values of temperatures and pressures given in Table 3.1, ρc has a value of 400 mks rayls. If the value of ambient temperature and barometric pressure differs appreciably from these values, then A_{cl} is determined using the Fig(3.5) [30].

Table 3.1:Ambient Pressure and Temperatures for Which ρc (Air)= 400 mks rayls

Ambient Pressure			Ambient Temperature T	
p_s N/m^2	m of Hg, 0 °C	in of Hg, 0 °C	0 °C	0 °F
0.7 x 10^5	0.525	20.68	-124.3	-192
0.8 x 10^5	0.600	23.63	-78.7	-110
0.9 x 10^5	0.675	26.58	-27.0	-17
1.0 x 10^5	0.750	29.54	30.7	87
1.013 x 10^5	0.760	29.9	38.9	102
1.1 x 10^5	0.825	32.5	94.5	202
1.2 x 10^5	0.900	35.4	164.4	328
1.3 x 10^5	0.975	38.4	240.4	465
1.4 x 10^5	1.050	41.3	322.4	613

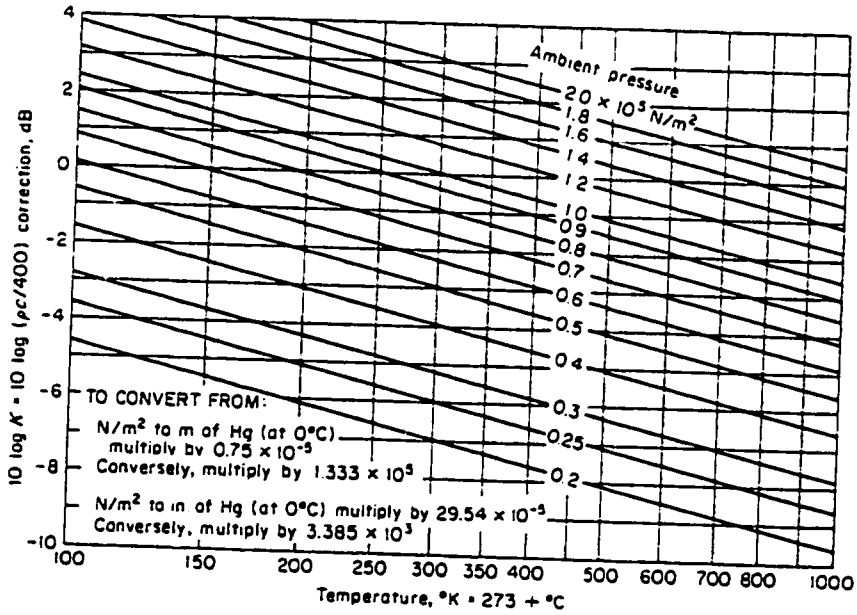


Fig.(3.5): Determining the value of 10 log K=10 log (ρc/400) as a function of ambient temperature and pressure

3.3.2 Excess Attenuation by Air Absorption

Two processes cause sound absorption in quite isotropic air: first, energy is extracted from a sound wave by losses arising from heat conduction and viscosity in the air. This is called classical absorption, proportional to the frequency square, and significant only at low temperatures. Second, energy is extracted from a sound wave by rotational and vibration relaxation of the oxygen molecules in the air. The vapor contents determine the time constant. Molecular absorption also depends upon temperature.

Atmospheric attenuation at a temperature of 20°C may be calculated from [30]

$$A_{e2} = 7.4\frac{f^2 r}{h_{\%}}10^{-8} \quad dB \tag{3.19}$$

where f = geometric-mean frequency of band (Hz)

$h_{\%}$ = % relative humidity

r = distance between source and receiver (m)

The geometric mean of the i^{th} octave band is assigned according to the following schedule [18]:

Table 3.2: Schedule for geometric mean of i^{th} octave

Octave	1	2	3	4	5	6	7	8	9	10	11
f_{min}	11	22	44	88	177	355	710	1420	2840	5680	11360
f_{max}	22	44	88	177	355	710	1420	2840	5680	11360	22720
f	16	31.5	63	125	250	500	1000	2000	4000	8000	16000

For other temperatures the following approximation can be used

$$A_{e2}(T,\phi=50\%)=\frac{A_{e2}(20^{\circ}C,\phi=50\%)}{1+\beta\Delta t\, f}\quad dB \tag{3.20}$$

where Δt = temperature difference in either °C of °F

β = constant ($4*10^{-6}$ for Δt in °C or $2*10^{-6}$ for Δt in °F)

Above equations are useful engineering estimates of A_{e2} atmospheric conditions near standard. More accurate information can be found in [31]. In our simulation we have used Eq.(3.19).

3.3.3 Excess Attenuation by Fog, Rain or Snow

It is said that on days of fog or light precipitation sound carries exceptionally well. During the light precipitation, the gradients of the temperature and wind tend to be small so that sound "carries" further outdoors than on a sunny day with the attendant micrometeorological inhomogeneities resulting from the sun's heating. It is stated [32] that molecular absorption due to fog is entirely negligible in the audible frequency range even for the dense fog.

3.3.4 Excess Attenuation Due to Ground Effect

The sound propagating above ground surface reaches the receiver without being reflected, i.e. the direct ray and the sound reflected by the ground. These rays will interfere at reception point. The amplitude and phase of the reflected sound ray will be different from the original one, and no complete cancellation will occur.

Generally, pronounced attenuation occurs in the frequency range 200-800 Hz. An increase of sound pressure occurs below 200 Hz and at frequencies above 800 Hz less attenuation is observed [33].

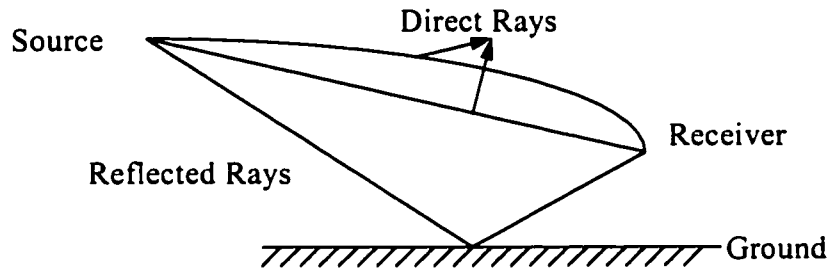


Fig.(3.6) Different paths of acoustic signal reaching the receiver

An empirical formula for sound attenuation due to ground effect is [33]:

$$A_{es} = \left[4.8 - \frac{2h_m}{r}\left(17 + \frac{300}{r}\right)\right] \tag{3.21}$$

with h_m being the average of source and receiver height, r is the distance between source and the receiver.

3.4 Flight Mission

In the simulation, we considered a helicopter flying at a constant height and speed. Flight path allows the simulation of Doppler shift, ground reflection, and effect of directivity as well. Helicopter can be considered coming from right to left or from left to right. By changing $x_0 \in [-850,+850]$,$y_0 \in [0,1000]$,$h_0 \in [50,600]$, and $u \in [0,60]$

(all the distances are in 'm' and velocity is in 'm/s'), a variety of flight scenarios can be simulated.

The Following figure shows a time domain helicopter sound signal and its spectrum generated by the use of simulator described in this chapter and its spectrum.

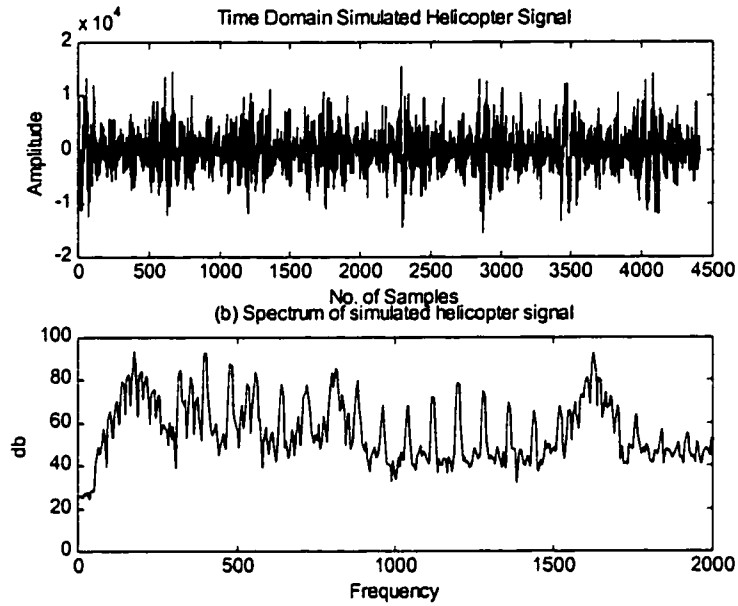


Fig.(3.7): (a) Time domain simulated helicopter signal (b) Spectrum of simulated helicopter signal

Chapter 4

Acoustic Signal Encoding

Summary

In designing a classifier, the choice of feature parameters of data is very important, since it greatly affects the overall performance of the system. The method of linear predictive analysis has been very successful for speech compression and has been utilized to obtain the feature vectors in speech recognition systems. The importance of the method lies both in its ability to provide an accurate estimate of acoustic parameters, and its computational efficiency. This chapter presents the formulation of the key ideas behind linear prediction. The chapter also reviews other alternatives to the LPC predictor coefficients such as, Reflection coefficients, LSP, and LPC cepstrum coefficients. These techniques will be used later to encode the signal into feature vectors in this thesis work.

4.1 Linear Predictive Coding of Acoustic Signal

Saito and Itakura [34] and Atal and Schroeder [35] proposed and applied linear predictive technique to speech analysis and synthesis in 1966. Since then, this technique has been proved to be very useful in providing an efficient representation of acoustic signal. The Linear Predictive Coding (LPC) model is based on the fact that a signal conveying a message is never completely random. Successive samples are correlated to each other. This correlation can be exploited to reduce the amount of data while keeping most of the information in the signal. So a parametric model of the behavior of the signal can be developed, which can later be used for prediction, control or data compression.

4.1.1 LPC Model

The signal s_n is considered to be wide sense stationary during a window or frame of length N. This acoustic sample, s_n, can be considered as the output of a system with some unknown input u_n using p previous output samples and q previous input samples as [36]

$$s_n = -\sum_{k=1}^{p} a_k s_{n-k} + G\sum_{l=0}^{q} b_l u_{n-l} \qquad \text{with } b_0 = 1 \qquad (4.1)$$

where G is a gain factor for input speech. Alternatively an all-pole model also called AutoRegressive (AR) model is achieved, by setting q=0 in the above equation [36], i.e.

$$s_n = -\sum_{k=1}^{p} a_k s_{n-k} + G u_n \tag{4.2}$$

In frequency domain, the above system can be represented as an all-pole transfer function as follows

$$H(z) = \frac{G}{1 + \sum_{k=1}^{p} a_k z^{-k}} \tag{4.3}$$

The transfer function in Eq.(4.3) can also be represented as [37]

$$H(z) = \frac{G}{A(z)} = \prod_{k=1}^{p} \frac{G}{1 - z_k z^{-1}} = \sum_{k=1}^{p} \frac{r_k}{1 - z_k z^{-1}} \tag{4.4}$$

where r_k is the residue of the pole z_k.

Helicopter acoustic signal is non-stationary, an accurate set of time invariant predictor coefficients can only be obtained over short intervals (typically 0.05 to 0.5 s) called frames, during which stationarity is assumed. An appropriate (10 or more) order linear prediction (LP) filter can represent short-term analysis filter. Before performing LP analysis, each frame of signal is first passed through a Hamming window. The LP analysis filter is given by

$$A(z) = 1 + \sum_{k=1}^{p} a_k z^{-k} \tag{4.5}$$

where p is the order of the LP filter.

The signal s_n can be predicted only approximately from a linearly weighted summation of past samples. Let this approximation of s_n be $\hat{s}_n$, where

$$\hat{s}_n = -\sum_{k=1}^{p} a_k s_{n-k} \tag{4.6}$$

Then the prediction error is given as

$$e_n = s_n - \hat{s}_n$$

$$e_n = s_n + \sum_{k=1}^{p} a_k s_{n-k} \tag{4.7}$$

a_k can then be obtained through minimization of the mean or total squared error with respect to each parameter. For deterministic signal we denote the total squared error as E, which is given as [36]

$$E = \sum e_n^2 = \sum_n \left(s_n + \sum_{k=1}^{p} a_k s_{n-k} \right)^2 \tag{4.8}$$

The minimization gives the following equation

$$\sum_{k=1}^{p} a_k \sum_n s_{n-k} s_{n-i} = -\sum_n s_n s_{n-i} \qquad \text{where } 1 \le i \le p \tag{4.9}$$

Eq.(4.9) form a set of p equations in p unknowns, which usually can be solved for predictor coefficients $\{a_k, 1 \le k \le p\}$. It is assumed that the error is minimized over the infinite duration $-\infty \le n \le +\infty$. So Eq.(4.9) reduces to

$$\sum_{k=1}^{p} a_k R(i-k) = -R(i) \tag{4.10}$$

where $\{1 \le k \le p\}$ and $R(i) = \sum_{n=-\infty}^{n=+\infty} s_n s_{n-i}$ is the auto-correlation function of signal s_n.

The coefficients $R(i-k)$ form an auto-correlation matrix, consequently this method is known as auto-correlation method. An auto-correlation matrix is a Toeplitz matrix, for which computationally efficient algorithm, known as Levinson-Durbin [38] recursion can be used. The Levinson-Durbin recursion is given in Appendix B.

4.2 PARCOR Coefficients (Reflection Coefficients)

Reflection coefficients are LP filter representations in an orthogonal system of coordinates. This makes these coefficients desirable in pattern matching task [39]. The value of the reflection coefficients does not change as the order of the filter is varied. The linear prediction coefficients obtained by the autocorrelation method [36] guarantee the stability of the all-pole filter, but due to use of window, the spectral resolution is reduced. In practice stability is not always guaranteed with finite word length (FWL) computation [39]. The problem can be resolved by using Itakura's [40] lattice formulation. This method guarantees filter stability with no windowing and relatively smaller sensitivity to FWL computation.

This method uses M forward and M backward linear predictors [40]. The m^{th} forward prediction error at time n is denoted by $x_m^+(n)$ and the m^{th} backward prediction error by $x_m^-(n)$. The forward prediction error, as shown in Fig.(4.1), results from difference between the signal and a predicted signal based upon a linear combination of samples at the previous m units of time [40]

$$x_m^+(n) = x(n) - \left[-\sum_{i=1}^{M} a_{mi} x(n-i) \right]$$

$$x_m^+(n) = \sum_{i=0}^{M} a_{mi} x(n-i) \qquad \text{with } a_{m0} = 1 \tag{4.11}$$

The backward prediction error results from the difference between the signal and a backward prediction based upon the following m samples, both evaluated at the time (n-m-1) so that only causal relations are involved.

$$x_m^-(n) = x(n-m-1) - \left[-\sum_{i=1}^{m} b_{mi} x(n-i) \right]$$

$$x_m^-(n) = \sum_{i=1}^{m+1} b_{mi} x(n-i) \qquad \text{with } b_{m,m+1} = 1 \tag{4.12}$$

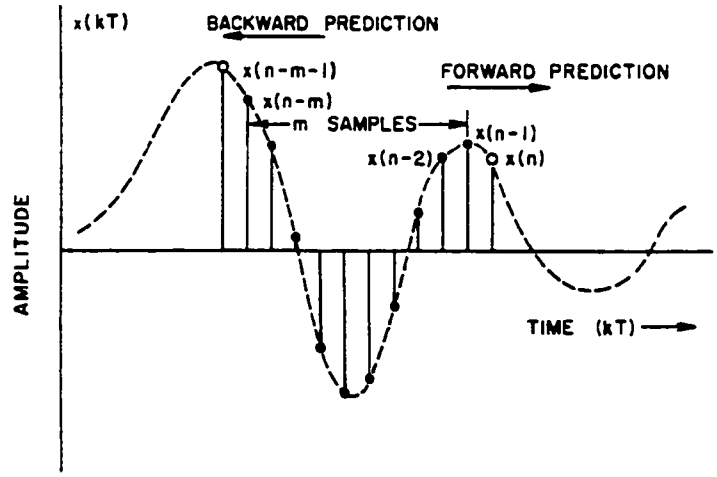


Fig.(4.1): Forward and backward prediction errors

The method consists of simultaneous minimization of the total square error of the forward and backward prediction errors α_m *and* β_m.

$$\alpha_m = \sum_{n=n_0}^{n_1} \left[x_m^+(n)\right]^2 \quad \text{and} \quad \beta_m = \sum_{n=n_0}^{n_1} \left[x_m^-(n)\right]^2$$

for m=1,2,.....,M

Taking partial derivatives with respect to specific parameters and equating them to zero carries out the minimization.

$$\frac{\partial \alpha_m}{\partial a_{mi}} = 0, \quad \frac{\partial \beta_m}{\partial b_{mi}} = 0 \quad i = 1,2,\cdots, M$$

The i^{th} stage LP analysis filter is given as [41]

$$A^{(i)}(z) = 1 + \sum_{k=1}^{i} a_k^{(i)} z^{-k} \tag{4.13}$$

The prediction error determined by Eq.(4.13) can be expressed in z-transform as

$$E^{(i)}(z) = A^{(i)}(z)\ S(z) \tag{4.14}$$

From Durbin recursion we know that [41]

$$a_k^{(i)} = a_{k-1}^{(i-1)} - k_i\, a_{i-k}^{(i-1)} \tag{4.15}$$

Taking the z-transform of Eq.(4.15) gives the following formula of $A^{(i)}(z)$ in terms of $A^{(i-1)}(z)$

$$A^{(i)}(z) = A^{(i-1)}(z) - k_i\, z^{-i}\, A^{(i-1)}\left(z^{-1}\right) \tag{4.16}$$

Putting Eq.(4.16) in Eq.(4.14) we will get

$$E^{(i)}(z) = A^{(i-1)}(z) S(z) - k_i\, z^{-i}\, A^{(i-1)}\left(z^{-1}\right) S(z) \tag{4.17}$$

In this expression, first term is the predictor error of $(i-1)^{th}$ order predictor. Define

$$B^{(i)}(z) = z^{-i}\, A^{(i)}\left(z^{-1}\right) S(z) \tag{4.18}$$

Now using Eq.(4.18), predictor error sequence (4.17) can be expressed as

$$E^{(i)}(z) = A^{(i-1)}(z) S(z) - k_i\, z^{-1}\, z^{-(i-1)}\, A^{(i-1)}\left(z^{-1}\right) S(z) \tag{4.19}$$

$$e^{(i)}(n) = e^{(i-1)}(n) - k_i\, b^{(i-1)}(n-1) \tag{4.20}$$

Substituting Eq.(4.18) in Eq.(4.20), we get

$$B^{(i)}(z) = z^{-i}\, A^{(i-1)}\left(z^{-1}\right) S(z) - k_i\, A^{(i-1)}(z) S(z) \tag{4.21}$$

$$E^{(i)}(z) = z^{-1}\, B^{(i-1)}(z) - k_i\, E^{(i-1)}(z) \tag{4.22}$$

So i^{th} stage backward prediction error is

$$b^{(i)}(n) = b^{(i-1)}(n-1) - k_i\, e^{(i-1)}(n) \tag{4.23}$$

Eq.(4.20) and Eq.(4.23) give the forward and backward prediction errors of i^{th} order predictor in terms of corresponding $(i-1)^{th}$ order predictor. PARCOR lattice is most easily visualized in terms of Z-transform. The analysis filter is shown in Fig.(4.2).

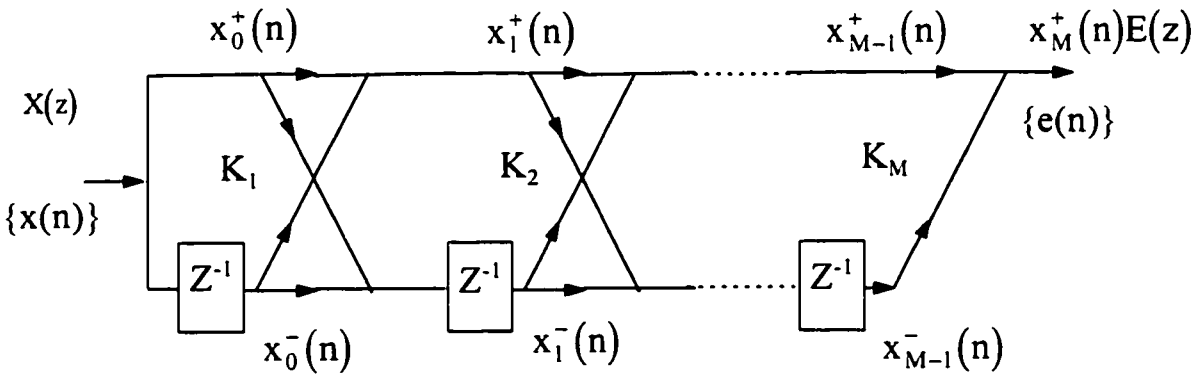


Fig.(4.2): PARCOR analysis filter

Due to the nature of the lattice structure the coefficients can be computed as [39]

$$k_i = \frac{\sum_{n=0}^{N-1} e^{(i-1)}(n)\, b^{(i-1)}(n-1)}{\left\{\sum_{n=0}^{N-1}\left(e^{(i-1)}(n)\right)^2 \sum_{n=0}^{N-1}\left(b^{(i-1)}(n-1)\right)^2\right\}^{\frac{1}{2}}} \tag{4.24}$$

Burg's algorithm [42] (Given in Appendix C) is used for efficient calculation of k_i.

4.3 Line Spectral Frequencies

Line spectral frequencies (LSFs) are an alternative parameterization of the filter with a one-to-one correspondence with the direct form predictor coefficients. The concept of LSFs ($0< \omega <\pi$) was first introduced by Itakura [43,44]. LSFs encode acoustic signal information more efficiently than other spectral parameters (log-area ratio, inverse sine transform) [45] due to intimate relationship between LSFs and the spectral peaks. Due to frequency domain interpretation, they provide frame to frame interpolation with smooth spectral changes. LSF parameters have well behaved dynamic range and are less susceptible to variations in acoustic characteristic and recording conditions. These LSFs in the cosine domain i.e., $x=\cos(\omega)$, are called line spectrum pairs (LSP).

The speech production model in terms of linear prediction coefficients a_k is given as

$$A(z) = 1 + \sum_{k=1}^{p} a_k z^{-k}$$

The time-reversed system function is given by [45]

$$B(z) = z^{-(p+1)} A(z^{-1}) \tag{4.25}$$

By expanding, the above two polynomials can be written as:

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{p-1} z^{-(p-1)} + a_p z^{-p} \tag{4.26}$$

$$B(z) = a_p z^{-1} + a_{p-1} z^{-2} + \cdots + a_2 z^{-(p-1)} + a_1 z^{-p} + z^{-(p+1)} \tag{4.27}$$

We can extend the order of the inverse filter $A(z)$ to (p+1) without introducing any new information by letting the $(p+1)^{th}$ coefficient to be 1 or -1. This is equal to setting the corresponding acoustic tube model completely closed or completely open at the $(p+1)^{th}$ stage [43]. Thus for ∓ 1, we respectively have

$$\begin{aligned} P(z) &= A(z) + B(z) \\ &= 1 + (a_1 + a_p) z^{-1} + \ldots + (a_p + a_1) z^{-p} + z^{-(p+1)} \\ &= 1 + \sum_{i=1}^{p+1} P_i z^{-i} \end{aligned} \tag{4.28}$$

with $P_i = a_i + a_{p+1-i}$ and $a_0 = 0$

$$\begin{aligned} Q(z) &= A(z) - B(z) \\ &= 1 + (a_1 - a_p) z^{-1} + \ldots + (a_p - a_1) z^{-p} - z^{-(p+1)} \\ &= 1 + \sum_{i=1}^{p+1} Q_i z^{-i} \end{aligned} \tag{4.29}$$

with $Q_i = a_i - a_{p+1-i}$ and $a_0 = 0$

The roots of these two auxiliary polynomials determine the LSFs.

By introducing a polynomial function all the zeros of P(z) and Q(z) can be made to lie on the unit circle. Since both P(z) and Q(z) have real coefficients only, their complex zeros occur in pairs. Therefore, if $e^{j\omega}$ is one of their zeros then $e^{-j\omega}$ is also one of their zeros. So for even p, P(z) and Q(z) can be written as [43]

$$P(z) = \left(1 + z^{-1}\right) \prod_{i=1,3,\ldots}^{p-1} \left(1 - 2\cos\omega_i z^{-1} + z^{-2}\right) \tag{4.30}$$

$$Q(z) = \left(1 - z^{-1}\right) \prod_{i=2,4,\ldots}^{p-1} \left(1 - 2\cos\omega_i z^{-1} + z^{-2}\right) \tag{4.31}$$

where the ω_i are real valued and correspond to the angles of the zero locations along the unit circle. The ω_i are called LSFs. They can be ordered such that

$$\omega_1 < \omega_3 < \cdots \omega_{p-1} < \omega_{p+1}$$

$$\omega_0 < \omega_2 < \cdots \omega_{p-2} < \omega_p$$

where $\omega_0 = 0$ and $\omega_{p+1} = \pi$. If the original LP polynomial has all its roots inside the unit circle, the LSP frequencies ω_i must be related by [43]

$$\omega_0 < \omega_1 < \omega_2 < \cdots < \omega_p < \omega_{p+1}$$

i.e. the roots of the two polynomials must be interlaced. This indicates that if the synthesis filter is stable, its coefficients can be represented as an ordered set of real numbers. Computation of LSP's is given in Appendix D.

4.4 Cepstrum Coefficients

Cepstrum analysis is another alternative for obtaining spectral feature vector. Cepstral analysis is widely used in the area of speech processing because it can represent speech waveform and its characteristics with fewer parameters. The cepstrum is a vehicle for separating much less informative pitch information from more important vocal tract shape information in speech processing. Cepstral parameters are usually obtained after LP analysis. It has been observed in speech recognition that better performance can be achieved using LPC cepstrum instead of LPC parameters as features [47,48,49]. Cepstrum coefficients have the advantage that frequency response distortions introduced by transmission system are removed [47].

As mentioned earlier, for a time series of samples s_n, a p^{th} order linear predictor $\hat{s}_n$, is a linear combination of p samples, denoted by

$$\hat{s}_n = -\sum_{k=1}^{p} a_k s_{n-k} \tag{4.32}$$

Let $S(e^{j\omega})$ be

$$S(e^{j\omega}) = \frac{\sigma^2}{1 + \sum_{k=1}^{p} a_k e^{-j\omega}} \tag{4.33}$$

LPC cepstrum is defined as the Fourier representation of the logarithmic amplitude spectrum based on LPC modeling. For a minimum phase LP filter the LPC cepstral coefficients c_n are defined as

$$log\left|S\left(e^{j\omega}\right)\right| = \sum_{n=0}^{\infty} c_n e^{-j\omega} \tag{4.34}$$

z-transform of above equation is

$$\log|S(z)| = \sum_{n=0}^{\infty} c_n z^{-n} \tag{4.35}$$

The predictor coefficients can be obtained using Partial Autocorrelation (PARCOR) algorithm [50]. A well-known recursive relation between predictor coefficients can be obtained by differentiating Eq.(4.35) with respect to z^{-1} and equating equal powers, yielding [48]

$$\begin{aligned} c_1 &= k_1 \\ c_n &= k_n + \sum_{m=1}^{n-1} \frac{m}{k} c_n \, k_{n-m} \end{aligned} \tag{4.36}$$

The use of the above recursion allows for an efficient computation of c_n and avoids polynomial factorization. Since c_n is of infinite duration, the feature vector of dimension p consists of the component c_1 to c_p, which are the most significant due to decay of sequence with increasing n.

Chapter 5

Wavelet Analysis

Summary

Wavelet transform (WT) is of interest for the analysis of non-stationary signals, because it provides an alternative to the short-term Fourier Transform (STFT) or Gabor transform. WT uses the notion of scale as an alternative to frequency, leading to a so-called time-scale representation (equivalent to time-frequency plane in STFT). This chapter provides a brief introduction to WT and focuses on signal processing application. The idea of multiresolution is applied for de-noising the helicopter acoustic signal. This de-noised version of the acoustic signal is used to get the feature vectors.

5.1 Introduction

Wavelet transform is a tool that splits data or functions or operators into different frequency components and then study these components with a resolution matched to its scale. WT is especially suited to non-stationary signal analysis. Wavelet representations provide a local time-frequency description of a signal. In the field of acoustic signal processing, they have been applied to analysis and synthesis of speech [51,52], formant extraction [53], noise suppression [54], speech compression [55], pitch detection [56], noise discrimination [57], speech coding [58] and acoustic speaker recognition [59].

The wavelet transform of a signal evolving in time (e.g., the amplitude of the pressure on an eardrum, for acoustic application) depends on two variables: scale (or frequency) and time, wavelets provide a tool for time-frequency localization.

5.2 Time Frequency Localization

The standard Fourier transform of a signal *f(t)* [60]

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int f(t)\, e^{-j\omega t}\, dt \tag{5.1}$$

gives a representation of the frequency contents of the signal *f(t)*, but information concerning time-localization of e.g., high frequency bursts cannot be read off easily from *F*.

Using windowed Fourier transform, by cutting off only a well-localized slice of *f(t)* and then taking its Fourier transform, time-localization is achieved.

$$\left(T^{WIN} f\right)(\omega, t) = \int f(s)\, g(s-t)\, e^{-j\omega s}\, ds \tag{5.2}$$

where *g(t)* is the window function.

For signal analysis, we usually use discrete version, where t and ω are assigned regularly spaced values $t = nt_0$ and $\omega = m\omega_0$, with $m, n \in N$, so

$$T_{m,n}^{WIN}(f) = \int f(s)\, g(s - nt_0)\, e^{-jm\omega_0 s}\, ds \tag{5.3}$$

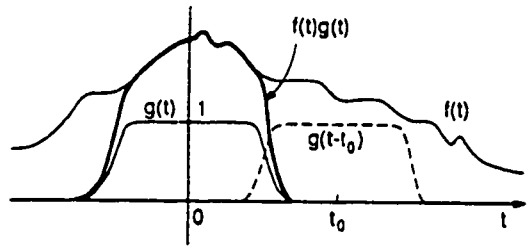


Fig.(5.1): Windowed Fourier transform

Fig.(5.1) above shows this procedure schematically. There are many choices for the window function g, most of which have compact support and reasonable smoothness.

5.3 The Wavelet Transform

The wavelet transform provides time-frequency description similar to windowed Fourier transform, with a few important differences. The wavelet transform formula analogous to (5.2) is [60]

$$(T^{WAV} f)(a,b) = |a|^{-\frac{1}{2}} \int f(t)\, \psi\left(\frac{t-b}{a}\right) dt \tag{5.4}$$

The function f can be constructed from its WT by means of "resolution of identity" formula [60] as

$$f = C_\psi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{da\, db}{a^2} \langle f, \psi^{a,b} \rangle \psi^{a,b} \tag{5.5}$$

where $\psi^{a,b}(x) = |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right)$, and < , > denotes the L^2-inner product. The constant C_ψ depends only on ψ.

$$C_\psi = 2\pi \int_{-\infty}^{+\infty} |\psi(\xi)|^2 \, |\xi|^{-1} \, d\xi \tag{5.6}$$

For discrete time WT formula analogous to (5.3) is [60]

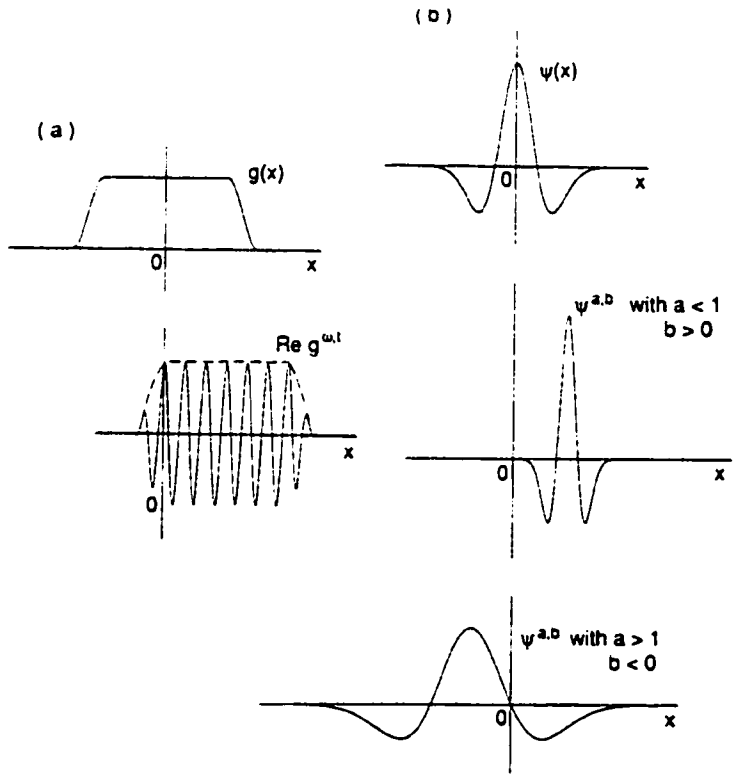
$$T_{a,b}^{WAV}(f) = |a|^{-\frac{1}{2}} \int f(t)\, \psi\left(a_0^{-m} t - n b_0\right) dt \tag{5.7}$$

with the assumption that $\int \psi(t)\, dt = 0$, a, the dilation parameter and b, the translation parameter, can attaining only discrete values:

$$a = a_0^m, \quad b = n b_0 a_0^m \quad with \quad m,n \in Z \quad and \quad a_0 > 1, b_0 > 0 \text{ fixed.}$$

Similarity between the wavelet and windowed Fourier transform is clear from formulas (5.2) and (5.4), which take the inner product of *f(t)* with a family of functions indexed by two labels, $g^{\omega,t}(s) = e^{j\omega s}\, g(s-t)$ in (5.2) and $\psi^{a,b}(s) = |a|^{-\frac{1}{2}} \psi\left(\frac{s-b}{a}\right)$ in (5.4). The function $\psi^{a,b}$ are called "wavelets", the function ψ is sometimes called "mother wavelet". As a changes, $\psi^{a,0}(s) = |a|^{-\frac{1}{2}} \psi\left(\frac{s}{a}\right)$

covers different frequency ranges (large values of the scaling parameter $|a|$ correspond to low frequencies, or coarse scale $\psi^{a,0}$, small values of $|a|$ correspond to high frequencies or very fine scale $\psi^{a,0}$). Changing the parameter b as well allows us to move the time localization center, each $\psi^{a,b}$ is localized around $s=b$. This follows that like (5.2), (5.4) also provides a time-frequency description of $f(t)$. The shapes of the analyzing function $g^{\omega,t}$ *and* $\psi^{a,b}$ differ in windowed Fourier transform and wavelet transform as shown in Fig.(5.2).



Fig(5.2): Typical shape of a) windowed Fourier transform function
b) wavelet function

In STFT, there is a restriction on time frequency resolution. Once a window, $g^{\omega,t}$, has chosen for the STFT, then the time-frequency resolution is fixed over the entire time-frequency plane. All the $g^{\omega,t}$, regardless of the value of ω have the same width, translated by proper time location. Contrary to this, the $\psi^{a,b}$ have time-widths adapted to their frequency: high frequency $\psi^{a,b}$ are very narrow, while low frequency $\psi^{a,b}$ are much broader. This gives wavelet transform a better capability to "zoom in" on very short-lived high frequency phenomena, such as transients in the signal.

5.4 Types of Wavelet Transforms

Starting with the equation (5.4) and (5.7) wavelet transform can be divided into two main types:

1. The continuous wavelet transform (Eq.(5.4))
2. The discrete wavelet transform (Eq.(5.7))

A very important class of wavelets in continuous and discrete case is called orthonormal basis of wavelets. They have made it possible to apply wavelet transform without redundancy.

5.4.1 The Continuous Wavelet Transform

Here the dilation and translation parameters a and b vary continuously over $\boldsymbol{R}$ (with the constraint $a \neq 0$), using formula (5.4). The continuous wavelet is the sum over all time of the signal multiplied by scaled and shifted versions of the wavelets. After

scaling a wavelet family, it is compared with the section at the start of the original signal. A number '*c*' is computed, which represents the correlation of the wavelet with the selected section of the signal as shown in Fig.(5.3 a) [61]. Higher '*c*' represents more similarity. Now shift the wavelet to right and find '*c*' for second portion as shown in Fig.(5.3 b) and so on till we have reached at the end of the signal. Increase the scale (stretch the wavelet) and again find '*c*' for this scale as shown in Fig.(5.3 c). Repeat this procedure for all scales. These coefficients, '*c*', given by Eq.(5.8) constitute the result of a regression of the original signal performed on the wavelets.

$$c_{a,b} = \int f(t)\, \psi^{a,b}(t) dt \tag{5.8}$$

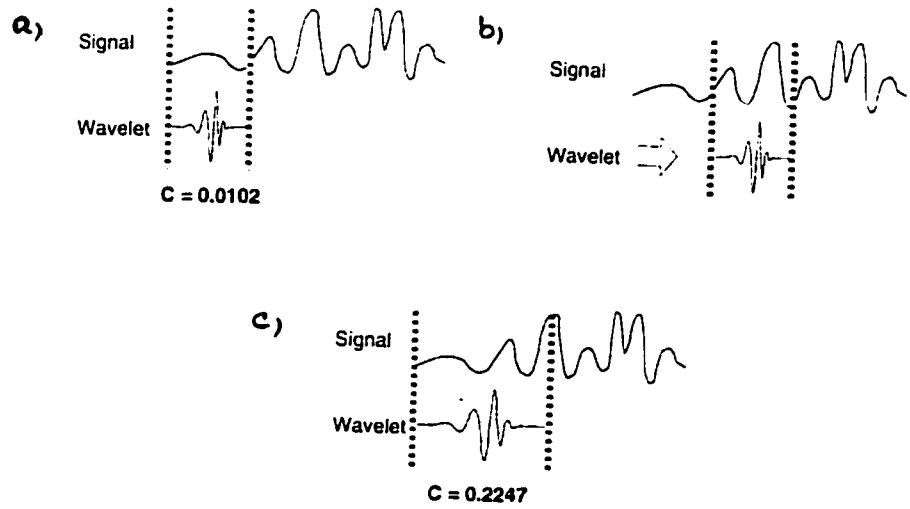


Fig.(5.3): Computation of wavelet coefficients

By drawing a plot on which the x-axis represents position along the signal (time), the y-axis represents scale (frequency), and the color at each x-y point represents the magnitude of the wavelet coefficients '*c*', the frequency contents of the signal can be analyzed as shown in Fig.(5.4).

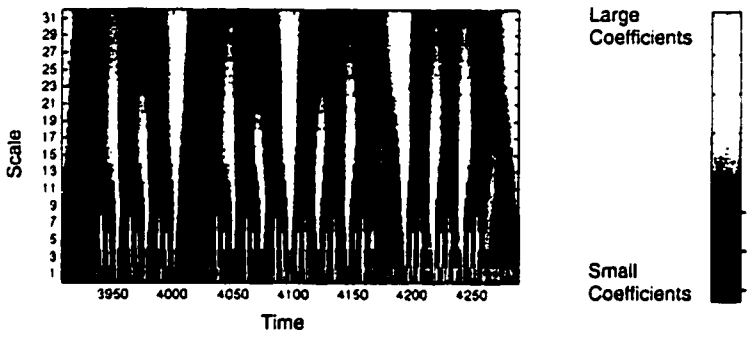


Fig.(5.4): Representation of wavelet coefficients

5.4.2 The Discrete Wavelet Transform

For discrete wavelet transform, the dilation parameter, *a*, and the translation parameter, *b*, both take only discrete values. For integer powers of one fixed dilation parameter $a_0 > 1$, $a = a_0^m$ is selected. As different values of *m* correspond to wavelets of different widths, the discretization of the translation parameter *b* depend on *m*. Narrow (high frequency) wavelets are translated by small steps in order to cover the whole time range, while wider (low frequency) wavelets are translated by large steps. The width of $\psi(a_0^{-m} x)$ is proportional to a_0^m; therefore to discretize *b* we choose $b = n b_0 a_0^m$, *where* $b_0 > 0$ is fixed and $n \in Z$. The corresponding discretely labeled wavelets are therefore [60]

$$\psi_{m,n}(x) = a_0^{-\frac{m}{2}}\ \psi\left(a_0^{-m}\left(x - nb_0a_0^m\right)\right) \tag{5.9}$$

$$= a_0^{-\frac{m}{2}}\ \psi\left(a_0^{-m}x - nb_0\right) \tag{5.10}$$

Like continuous wavelet transform, these discrete wavelet transforms provide a very redundant description of the original signal. This redundancy can be exploited or eliminated to reduce the transform to its bare essentials.

5.4.3 Orthonormal Wavelet Basis

If a is close to 1, then the wavelet functions are over complete and signal reconstruction takes place within non-restrictive conditions on ψ. Contrary to this, if the sampling is sparse, e.g. computation is done octave by octave (a=2) then we might miss some important frequency information. If a tight frame is such that all the wavelets are necessary to reconstruct a general signal, then the wavelet form an orthonormal basis. Only a special choice of ψ will give orthonormal basis. An arbitrary signal can then be represented exactly as a weighted sum of basis functions.

$$f(t) = \sum_{j,k} c_{a,b}\psi^{a,b}(t) \tag{5.11}$$

The basis function $\psi^{a,b}(t)$ are not only obtained from a single "mother wavelet" by means of scaling and shift but are also required to form an orthonormal basis. Examples of orthonormal basis wavelets for continuous and discrete wavelet transform are Daubechies, Haar, Coiflets, Symlets etc. [61]. We have used

Daubechies wavelet in our analysis. A few members of Daubechies wavelet family are shown in Fig.(5.5) below.

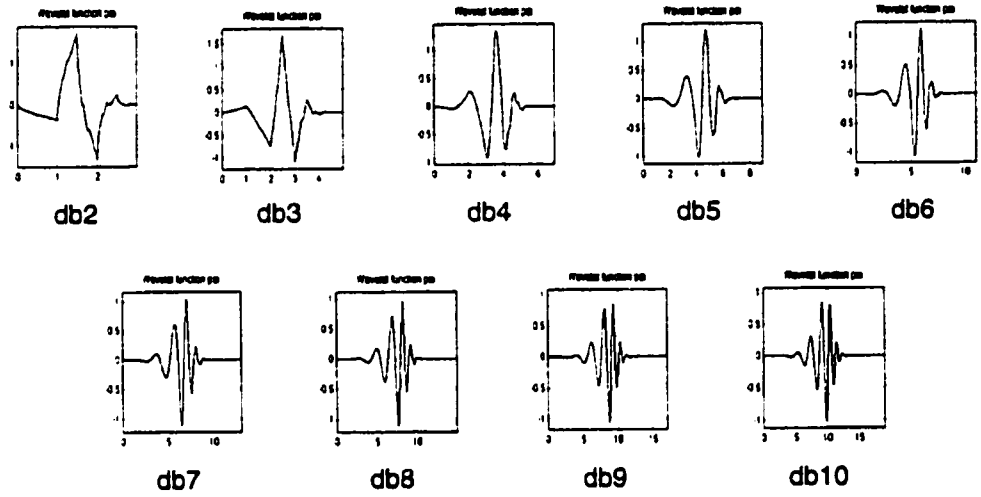


Fig.(5.5): Daubechies orthonormal wavelets

5.4.3.1 Orthonormal Basis in Discrete Time Case

To get an orthonormal basis in the discrete time case, two methods were developed which led to discrete wavelet transform.

1. Pyramidal coding or multiresolution signal
2. Subband coding

5.4.3.1.1 Multiresolution Pyramid

The lower resolution of signal, *x(n),* is obtained by low pass filtering with a half band low-pass filter with impulse response *g(n)* [12]. Due to reduction in frequency band, following Nyquist's rule, we can subsample by two (keeping alternate samples), thus doubling the scale in the analysis. The resultant signal *y(n)* is

$$y(n) = \sum_{k=-\infty}^{+\infty} g(k)\, x(2n-k) \tag{5.12}$$

The loss of high frequency components will change the resolution.

Now the approximation *a(n)*, of *x(n)* can be obtained from the low-passed and subsampled version *y(n)*. First we upsample *y(n)* by two (inserting zeros between the samples) since we need a signal at the original scale for comparison.

$$y'(2n) = y(n) \quad and \quad y'(2n+1) = 0$$

Then $y'(n)$ is interpolated with a filter with impulse response $g'(n)$ to obtain approximation *a(n)*.

$$a(n) = \sum_{k=-\infty}^{+\infty} g'(k)\, y'(n-k) \tag{5.13}$$

If *g(n)* and $g'(n)$ were perfect half band filters, then *a(n)* would be a perfect half band low pass approximation to *x(n)*. The high frequency details *d(n)* of the signal can be computed by the difference of original signal, *x(n)*, and half band approximation *a(n)*.

$$d(n)=x(n)-a(n) \tag{5.14}$$

We can construct *x(n)* by adding *a(n)* and *d(n)* as shown in Fig.(5.6).

There is some redundancy, since a signal with sampling rate f_S is mapped into two signals *d(n)* and *y(n)* with sample rate f_S and $f_S/2$ respectively. But due to perfect half band characteristics of the low pass filter, *d(n)* can be subsampled by two without loss of information.

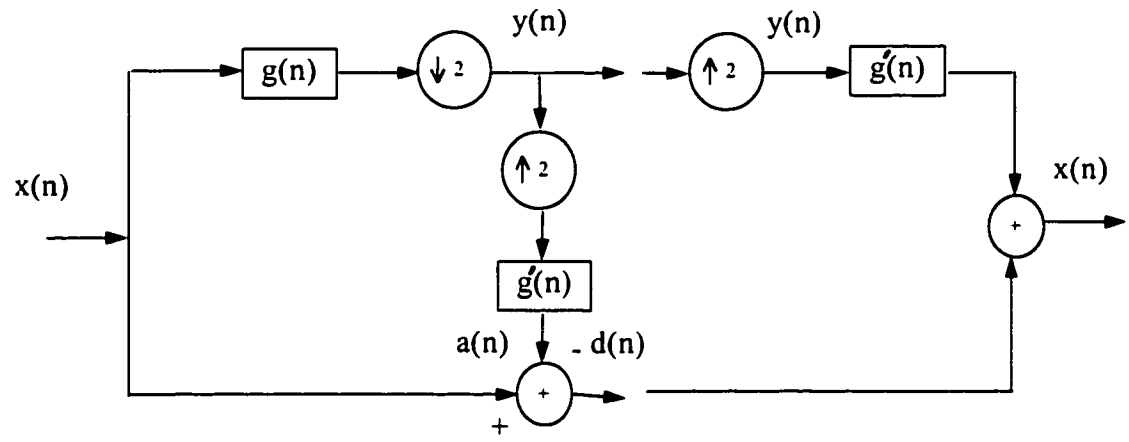


Fig.(5.6): Pyramid scheme to get lower band approximation a(n) and reconstruction of the original signal

Because of the resolution change due to low pass filtering and subsampling, this method is called Multiresolution Signal Analysis [63]. Same scheme can be iterated on *y(n)*, creating hierarchy of low-resolution signals. Due to hierarchy and sample reduction such schemes are signal pyramids.

5.4.3.1.2 Subband Coding Scheme

Pyramid resolution leads to a half rate low-resolution signal and a full rate difference signal, resulting in an increase in number of samples by 50%. In subband coding scheme no such redundancy appears [63]. In it, low pass subsampled approximation is obtained exactly like pyramid resolution, but instead of a difference signal, we compute the "added details" as a high pass filtered version of *x(n)*. So we use an ideal low pass and an ideal high pass filter. Proceeding in this manner discrete version is identical to continuous wavelet transform.

x(n) can be recovered from its two filtered and subsampled versions $y_0(n)$ and $y_1(n)$ without the use of ideal filters. Both $y_0(n)$ and $y_1(n)$ are upsampled by two and filtered with $g'(n)$ and $h'(n)$ and added together as shown in Fig.(5.7).

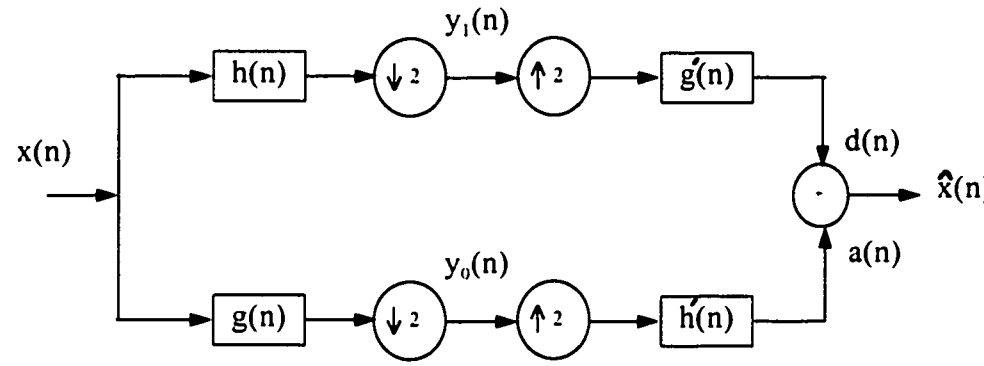


Fig.(5.7): Two subband approximations and the constructed signal

If we have filters with perfect reconstruction property then reconstructed signal will be identical to original signal *x(n)*. If we have identical (within time reversal) analysis and synthesis filters with perfect reconstruction property then this corresponds to orthonormal basis decomposition. The reconstruction is obtained by summing up orthonormal projection. Let the FIR high pass and low pass filters be related by

$$h(L-1-n)=(-1)^n g(n) \tag{5.15}$$

where L is filter length.

Applying the filter bank theory we can define such perfect reconstruction filter banks [64,65].

By means of orthogonal filters, a sequence *x(n)* can be decomposed into two half rate or half resolution sequences. This process can be iterated on either or on both sequences. To get the finer frequency resolution, we iterate the scheme on lower band. If *g(n)* and *h(n)* are good low pass and high pass filters respectively then the iteration of the scheme on the first low band creates a new low band correspond to lower quarter of the frequency spectrum. Each further iteration halves the width of the low band and increases the frequency resolution by two. At each iteration, the current high pass portion corresponds to the difference between the previous and the current low band portions. This is equivalent to Fig.(5.8) and the frequency resolution is shown in Fig.(5.9). The complexity involved in discrete wavelet transform is relatively low.

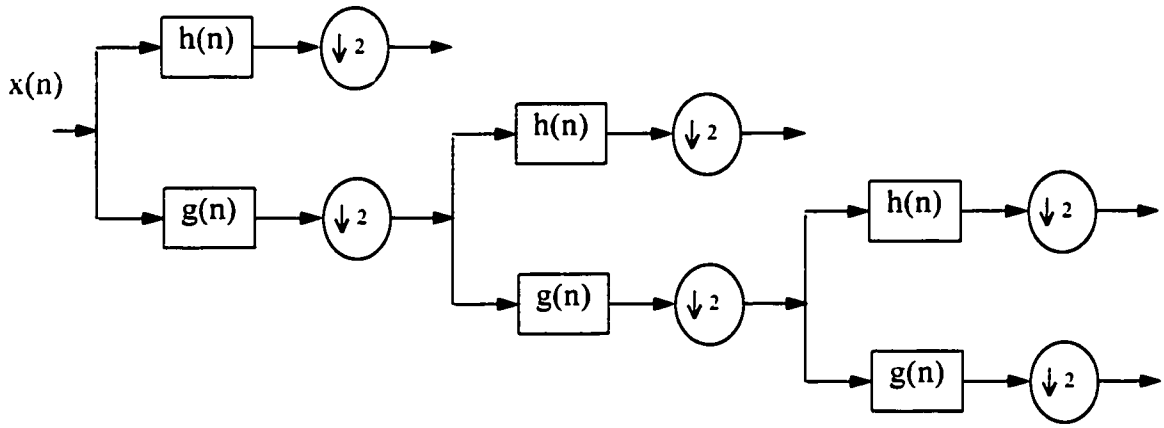


Fig.(5.8): Block diagram of discrete wavelet transform using discrete time filters

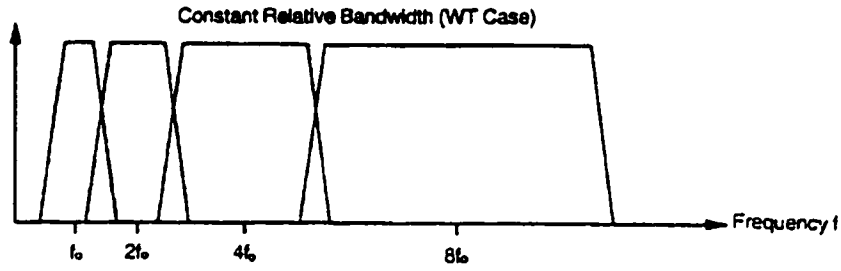


Fig.(5.9): Division of the frequency domain

5.5 De-noising of Signal

To remove noise from a signal using wavelets, we have to identify which component or components contain the noise and then reconstruct the signal without those components. By successive approximations, signal becomes less and less noisy as more and more high frequency information is filtered out of the signal. In discarding all the high frequency information, we may loose many of the sharp features of the original signal. To perform optimal de-noising a better approach called "thresholding" can be used. This involves discarding those portions of the details which disagree a certain limit. Entropy based criteria can be adopted for most suitable selection of the de-noising threshold. Threshold entropy is defined as

$$\begin{aligned} E(x_i) &= 1 \quad if\,|x_i| > \varepsilon \\ &= 0 \quad otherwise \end{aligned} \tag{5.16}$$

i.e. $E(x_i)$ is equal to a number when the signal is greater than a threshold p. Wavelet Toolbox [61] command "DDECCMP" is used to determine the threshold and the "WDENCMP" will de-noise the signal using this threshold.

Chapter 6

Artificial Neural Networks

Summary

Multilayer perceptrons have been successfully applied for the solution of difficult non-linear problems. They have the ability to learn from experience through training. The training of ANN can be carried out using the highly popular back-propagation algorithm. The algorithm is based on error correction learning rule. For a classifier design using ANN, the ultimate objective is to achieve an acceptable rate of classification through training. This chapter gives a brief history of ANN with some preliminaries to pave the way for the derivation of back-propagation algorithm. The reasons for using ANN in signal processing area are given and are followed by the derivation of back-propagation algorithm.

6.1 History

Modern era of neural networks begun with the pioneering work of McCulloch and Pitts [66] in 1943, in which they have described a logical calculus of neural networks. In 1956 Uttley [67] demonstrated that neural networks with modified synapses might learn to classify binary patterns. Rosenblatt [68] proposed the perceptron convergence theorem in 1958 and proved his theorem in 1960. In 1969 Minsky and Papert [69] demonstrated that there is a fundamental limit to one-layer perceptrons computation. They further stated that any of the virtues of the one-layer perceptrons carry over to the many-layered version. Werbos [70] in 1974 developed the back propagation algorithm to propagate the error at output back to the hidden layer. The most influential publications are the 1982 paper by Hopfield [71] and the two-volume book by Rumelhart and McLelland [71], which are responsible for the resurgence of interest in neural networks in the 1980s. Since then there is an exponential rise in the research contributions to this field.

6.2 Architecture

Artificial neural networks are biologically inspired information-processing units. They consist of simple processing elements called neurons. These neurons have the ability to work in parallel. These neurons are interconnected and the strength of interconnection is denoted by the parameter called synaptic weights. The model of a

neuron is shown in Fig.(6.1), $x_1, x_2, \ldots, x_p$ are the input signals, $w_{k1}, w_{k2}, \ldots, w_{kp}$ are the synaptic weights of neuron k, w_{k0} is the bias, v_k is the linear combiner output, $\varphi(.)$ is the activation function, and y_k is the output signal of the neuron. In mathematical terms the neuron k is described as [72]

$$v_k = \sum_{j=1}^{p} w_{kj} x_j + w_{k0}$$
$$y_k = \varphi(v_k) \tag{6.1}$$

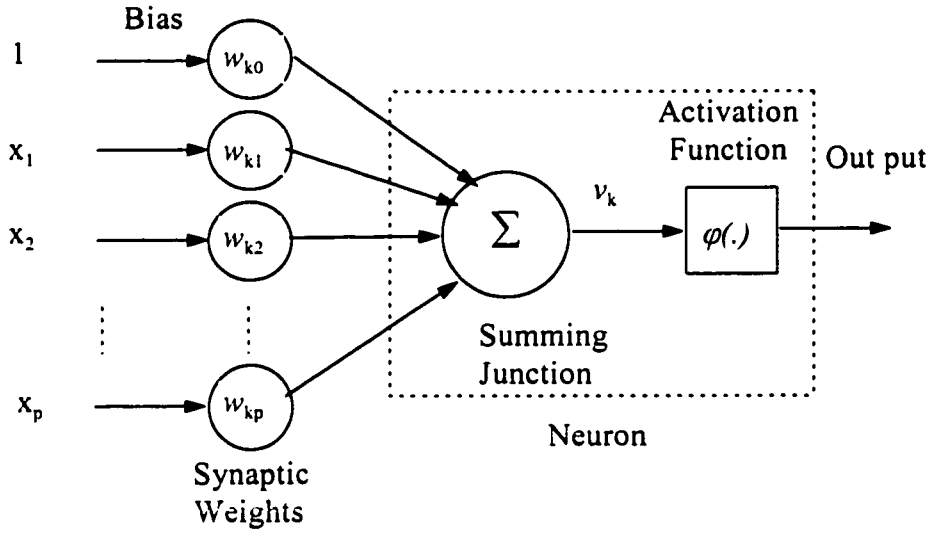


Fig.(6.1): Model of a neuron

6.3 Activation Function

The activation function, denoted by $\varphi(\cdot)$, defines the output of a neuron in terms of the activity level at its input. Various types of non-linear activation functions can be

used. Some of them are shown in the Fig.(6.2). For the problem under study sigmoid or hyperbolic tangent sigmoid function can be used in the hidden layers and any other function can be used for the output layer. The hyperbolic tangent sigmoid function is defined as:

$$\varphi(v_k) = \frac{1 - \exp(-av_k)}{1 + \exp(-av_k)} \tag{6.2}$$

where a is a constant. In our application a is 1.

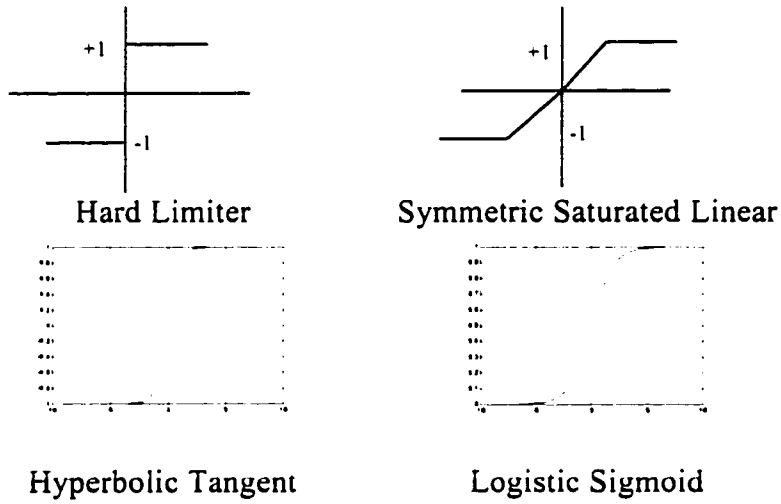


Fig.(6.2): Types of activation function

6.4 Multilayer Feed-forward Networks

Multilayer feed-forward networks consists of an input layer, a number of hidden layers and an output layer as shown in Fig.(6.3) [73]. It is the most popular neural network structure. Input layer is essentially a direct link to the inputs of the first hidden layer. The hidden layer consists of neurons with non-linear activation function

in our application. The output neuron in our case has a linear activation function. The output of each neuron is connected to the inputs of all the neurons in the next layer. Signal is unidirectional i.e., flows only from input to output.

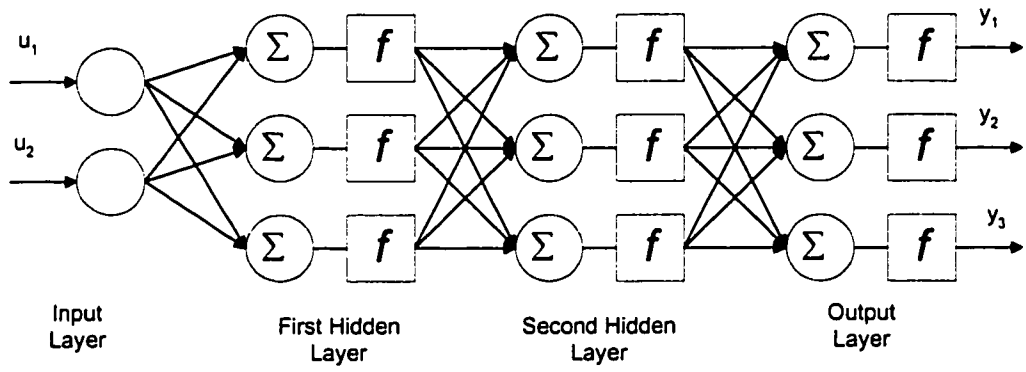


Fig.(6.3): Multi-layer feedforward NN

6.5 Neural Network and Signal processing

Methods for statistical signal processing are founded on three assumptions: linearity, stationarity and second order statistics with particular emphasis on Gaussianity. These assumptions are made for the sake of mathematical tractability. Most, if not all, physical signals in real life applications are non-linear, non-stationary and non-Gaussian, resulting in a sub-optimal signal processor. The performance of these systems can be improved by the use of neural networks in combination with suitable feature extraction techniques (e.g., time-frequency analysis) [74,75].

Following are a few characteristics that befit the use of neural networks for signal processing applications [76]:

- Each neuron (processing unit) in neural network has a non-linear built-in activation function that enables the modeling of the underlying non-linearities in the physical mechanism.
- In a multilayer neural network, each neuron in a particular layer is connected to a large number of source/node neurons in the previous layer. This form of global interconnectivity has the potential to fault tolerant. If a neuron or its synaptic links are damaged, the quality of classification is impaired, but due to the highly distributive nature, the damage has to be extensive before the performance is seriously degraded.
- Neural networks have a natural ability to adapt their free parameters to statistical changes in the environment in which they operate.
- Neural networks undergo training session in which their free parameters (synaptic weights and biases) are adjusted in a systematic way to minimize the mean-square-error between the desired response and actual output of the network in response to the input signal. So a "non-parametric" approach for estimation is applied, meaning no knowledge of underlying probability distribution is required.
- For sufficient number of hidden units, neural networks can approximate any continuous input-output mapping to any desired degree of approximation. This property makes the neural networks as Universal Approximators [73,77].

To capture the full information contents of the input data in a computationally efficient manner, neural networks may have to be integrated with some signal

transformation technique (such as linear prediction coefficients, reflection coefficients, line spectral pairs, cepstrum coefficients etc. in the proposed application). Whatever system integration is used, the available information pertaining to a signal processing task (target detection in this case) should be preserved optimally (in a statistical sense) and used efficiently (in a computational sense), until the receiver is ready for final decision-making.

6.6 Training of Feed-forward Neural Networks

Typical training of a multilayer feed-forward network using the back-propagation algorithm can be described as follow [73]:

Starting with a suitable network configuration, set all the synaptic weights and biases of the network to small uniformly distributed random numbers. The set of the input data and the desired outputs are collected. The input set of data is applied to the neural networks and the output of the network is compared with desired outputs. The error so generated is used to train the feed-forward network through back-propagation of this error. The synaptic weights are adjusted while minimizing a cost function, typically sum of the squared error. The synaptic weights are adjusted iteratively until a stopping criterion is met.

Nomenclature:

$w_{ji}(.)$ *weight connecting the output of neuron i to the input of neuron j*

$\Delta w_{ji}(.)$ *weight correction factor.*

$e_j(.)$ *error signal at the output of neuron j*

$\varepsilon(.)$ *instantaneous sum of the square error*

ε_{av} *average squared error*

$d_j(.)$ *desired response of neuron j*

$y_j(.)$ *output of neuron j*

$\varphi_j(.)$ *activity function*

6.6.1 Derivation of Back-propagation Algorithm:

Back-propagation is a first order gradient scheme to minimize the error between the actual and the desired output. The error signal at the output of neuron j at iteration n as shown in Fig.(6.4) is defined by

$$e_j(n) = d_j(n) - y_j(n) \tag{6.3}$$

Let the instantaneous squared error for neuron j be $\frac{1}{2}e_j^2(n)$. The instantaneous sum of squared error of the network is given as

$$\varepsilon(n) = \frac{1}{2}\sum_{j\in C} e_j^2(n) \tag{6.4}$$

where C includes all the neurons in the output layer of net network.

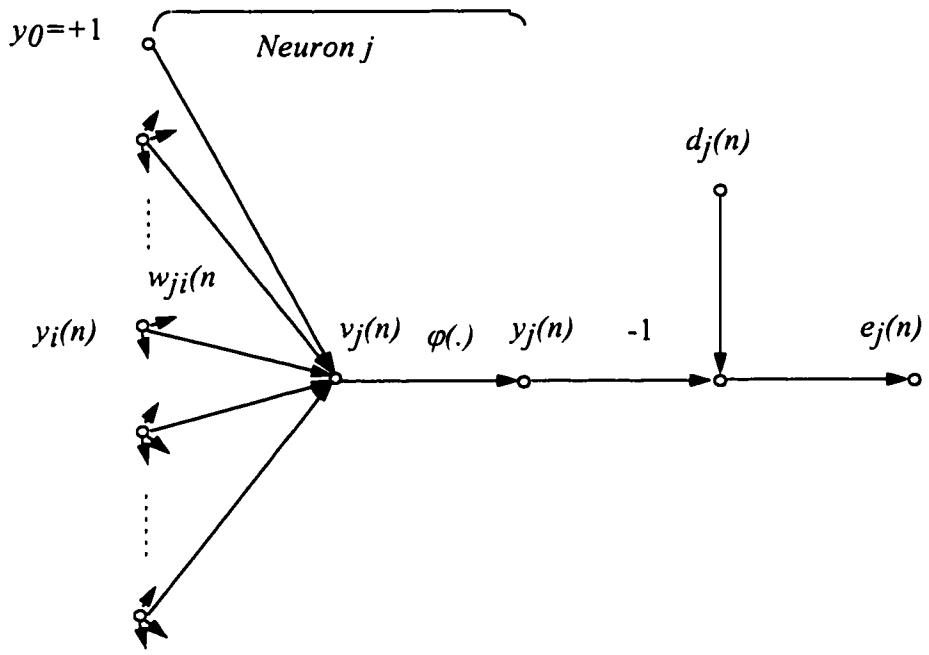


Fig.(6.4): Signal flow graph showing the details of output neuron j

If N denotes the total number of patterns in the input, then the average squared error is given as

$$\varepsilon_{av} = \frac{1}{N}\sum_{n=1}^{N}\varepsilon(n) \tag{6.5}$$

The average squared error ε_{av}, is a function of all the free parameters (synaptic weights and biases). The objective of the learning process is to adjust the free parameters of the network so as to minimize ε_{av}.

The net internal activity level $v_j(n)$ produced at the input of the nonlinearity associated with the *jth* neuron is

$$v_j(n) = \sum_{i=0}^{p} w_{ji}(n)y_i(n) \tag{6.6}$$

where p is the total number of inputs.

$$y_j(n) = \varphi_j\left(v_j(n)\right) \tag{6.7}$$

The corrections $\Delta w_{ji}(n)$ to the synaptic weights w_{ji}, are proportional to the instantaneous gradient $\partial\varepsilon(n)/\partial w_{ji}(n)$. According to chain rule

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} = \frac{\partial \varepsilon(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \tag{6.8}$$

Differentiating Eq.(6.3), Eq.(6.4), Eq.(6.6) and Eq.(6.7) with respect to the appropriate variable give the following relation

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} = -e_j(n)\varphi_j'\left(v_j(n)\right)y_i(n) \tag{6.9}$$

The corrections $\Delta w_{ji}(n)$ to the synaptic weights $w_{ji}(n)$ is defined by the delta rule:

$$\Delta w_{ji}(n) = -\eta \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} \tag{6.10}$$

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \tag{6.11}$$

where $\delta_j(n)$ is called local gradient defined as

$$\delta_j(n) = -\frac{\partial \varepsilon(n)}{\partial e_j(n)} \cdot \frac{\partial e_j(n)}{\partial y_j(n)} \cdot \frac{\partial y_j(n)}{\partial v_j(n)} \tag{6.12}$$

$$= e_j(n)\, \varphi_j'\left(v_j(n)\right) \tag{6.13}$$

Above equation shows that the key factor involved in the weight adjustment $\Delta w_{ji}(n)$ is the error signal at the output of the neuron. In this context, there are two distinct cases depending upon where neuron j is located:

Case I: Neuron j is an output node

Case II: Neuron j is a hidden node

6.6.1.1 Case I: Neuron j is An Output Node

When neuron j is in the output layer, we have the desired response of that neuron. So we can determine the error using Eq.(6.3) and can find the local gradient using Eq.(6.12).

6.6.1.2 Case II: Neuron j is A Hidden Node

When neuron j is located in the hidden layer, there is no specific desired response of the neuron. But here again it shares the responsibility for any error made at the output of the network. Consider the situation shown in Fig.(6.5) depicting j as a hidden neuron [74].

Let the local gradient is redefined as

$$\delta_j(n) = -\frac{\partial \varepsilon(n)}{\partial y_j(n)} \cdot \frac{\partial y_j(n)}{\partial v_j(n)} \tag{6.14}$$

$$= \frac{\partial \varepsilon(n)}{\partial y_j(n)} \varphi_j'\left(v_j(n)\right) \tag{6.15}$$

The instantaneous sum of squared error of the network is given as

$$\varepsilon(n) = \frac{1}{2}\sum_{k \in C} e_k^2(n) \qquad \text{neuron k is an output node} \tag{6.16}$$

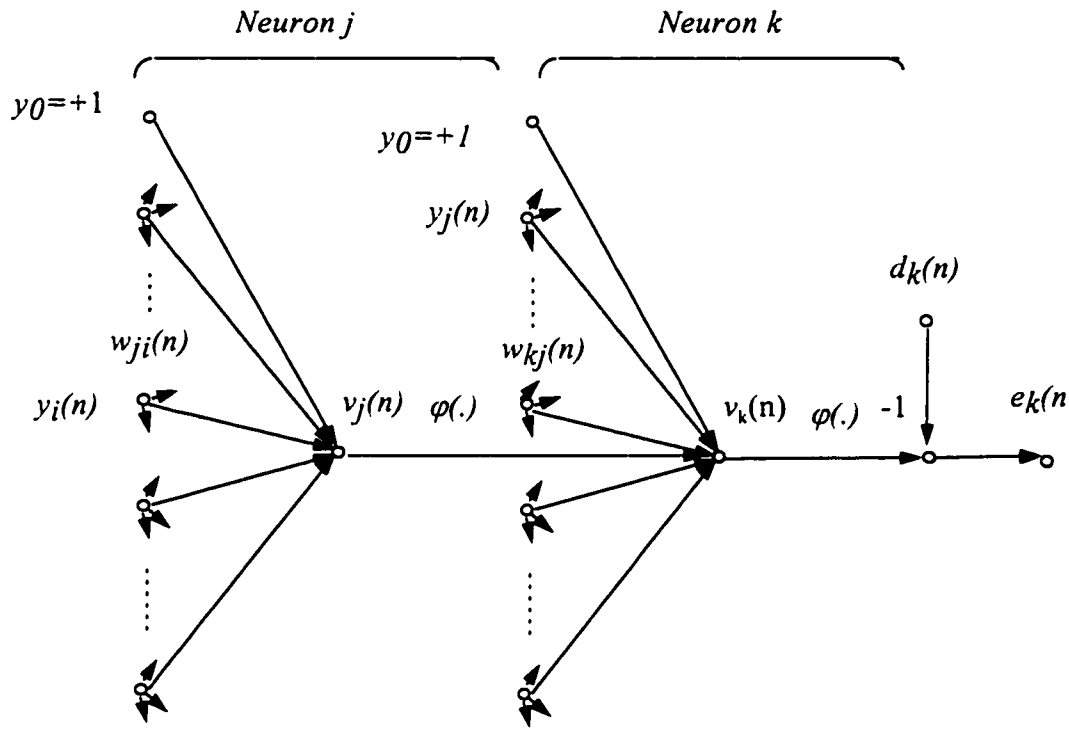


Fig.6.5: Signal flow graph showing the details of output neuron k and hidden neuron j

The partial derivative of the above equation with respect to $y_j(n)$ is

$$\frac{\partial \varepsilon(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial y_j(n)} \tag{6.17}$$

Using chain rule above equation can be written as

$$\frac{\partial \varepsilon(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial v_k(n)} \cdot \frac{\partial v_k(n)}{\partial y_j(n)} \tag{6.18}$$

As neuron k is in the output layer so we have access to its error

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -\varphi_k'\left(v_k(n)\right) \tag{6.19}$$

From Fig.(6.5) it is also clear that activity at neuron k is

$$v_k(n) = \sum_{j=0}^{q} w_{kj}\, y_j(n) \tag{6.20}$$

q is the total number of inputs to neuron k.

which will give

$$\frac{\partial v_k(n)}{\partial y_j(n)} = w_{kj} \tag{6.21}$$

Putting Eq.(6.19) and Eq.(6.21) in Eq.(6.18) we have

$$\frac{\partial \varepsilon(n)}{\partial y_j(n)} = \sum_k \delta_k(n)\, w_{kj}(n) \tag{6.22}$$

Now putting Eq.(6.22) in Eq.(6.15) local gradient for neuron j results as

$$\delta_j(n) = \varphi_j'\left(v_j(n)\right) \sum_k \delta_k(n)\, w_{kj}(n) \tag{6.23}$$

i.e. the local gradient is the product of the associated derivative $\varphi_j'\left(v_j(n)\right)$ and the weighted sum of δ's for the neuron in the next hidden or output layer.

The back-propagation algorithm can be expressed as:

$$\begin{pmatrix} \textit{Weights} \\ \textit{correction} \\ \Delta w_{ji}(n) \end{pmatrix} = \begin{pmatrix} \textit{Learning} \\ \textit{rate parameter} \\ \eta \end{pmatrix} . \begin{pmatrix} \textit{Local} \\ \textit{gradient} \\ \delta_j(n) \end{pmatrix} . \begin{pmatrix} \textit{input signal} \\ \textit{of neuron } j \\ y_i(n) \end{pmatrix}$$

6.6.1.3 Stopping Criteria

The back-propagation algorithm is considered to have converged when the cost function or sum square error (SSE) measure of the network has reached a pre-determined value, say 0.1 or the change in the SSE is almost constant.

In short the backpropagation algorithm for training of ANN is as follow

1. network initialization
2. presentation of training examples
3. forward pass computation
4. backward computation to determine correction factor
5. correction in synaptic weights obtained by backward computation
6. go to 4 till stopping criteria is met

Chapter 7

Helicopter Detection: Simulation Studies

Summary

In this chapter some simulation results are presented. Different LPC parametric modeling approaches (resulting in different feature vectors) were applied to ANN based detection system. The performance of these features on the recognition results is presented. As the performance has been degraded with the increase of background noise, wavelet transforms were applied for de-noising the signal prior to obtaining the LPC features. This resulted in improvement in the detection. Finally a system was trained using a combination of clean as well as noisy frames of signal. Effect of this training strategy on the overall detection is presented.

7.1 Feature Extraction

In this study, the ANN detection system was trained on helicopter and non-helicopter sounds. Four different helicopters S-67, H500, CH-47C and MI-HIND were considered in this problem. Helicopters were assumed to pass by a stationary observer. The helicopter sounds were obtained using the simulator described in chapter 3. With a specific speed, starting point and definite height, a 28 seconds simulated sound for each helicopter was obtained. The sound was sampled at 11kHz using 16 bits per sample. Similarly musical sound, English and Arabic conversations were also recorded at 11kHz. These recordings are of different time duration (6-second recordings for music and English conversation and two 10-second recordings for Arabic conversations).

The length of the recordings refers to the amount of the usable data. These helicopter and non-helicopter sound recordings were split into segments of uniform length, called frames, to produce training and test patterns. Each frame was of 0.1 second duration i.e.; it contained around 1102 data points. It was assumed that spectral characteristics remain constant during the frame length (i.e., signal is wide sense stationary). To get the training set of patterns we have chosen every fourth frame leaving three between the successive selections (i.e., we have selected 1^{st}, 5^{th}, 9^{th}, . . . frames to produce training set of data from the helicopter signal). Alternate frames of non-helicopter sound frames were included in the training set as the length of non-helicopter data was short. To produce test data we have selected 3^{rd}, 7^{th}, 11^{th}, .

. . frames from helicopter data bank and 2^{nd}, 4^{th}, 6^{th}, . . frames for non-helicopter sound recordings.

The analysis of helicopter sound in Chapter 3 has shown that the main characteristics of helicopter acoustic signal lies below 2kHz. In order to remove the irrelevant information we have used a low pass filter with 2kHz cut off frequency. Digital Elliptic filter is applied due to steeper rolloff characteristics as compared to the same order Butterworth or Chebyshev filters. A 5^{th} order filter, with 3db of ripple in the passband and a stopband 30db down from the peak value in the passband, is used. Further the signal was sampled at 5kHz. We have applied an anti-aliasing FIR filter to limit the highest frequency in the acoustic signal to 2kHz. The Kaiser window is used for designing this FIR digital filter. Then the signal was smoothed using a Hamming window. From this lowpass version of the acoustic signal the feature vectors were obtained using different acoustic encoding techniques discussed in Chapter 4. These steps are shown in Fig.(1.4).

ANN identifiers of different sizes were trained using these feature vectors and their performance was measured in terms of the their classification ability. The results for different encoding techniques are discussed in the following sections.

7.2 Detection Using the LPC Coefficients as Features

An ANN recognition system uses a set of features to distinguish between helicopter and non-helicopter acoustic signals. Features are characteristic attributes calculated

to discriminate between the classes of patterns. Search of invariant or useful features is very important in an identification and detection problem. Spectral modeling technique through LPC, which is very popular in speech recognition, was used to get the feature vectors. It is a parametric modeling approach i.e., the spectrum is completely represented by a relatively small set of parameters. For the purpose of feature extraction, parametric modeling representation offers a richer variety of possibilities by allowing the flexibility of applying various mathematical operations on the model. Parameterization can be considered as a process of redundancy removal.

We have assumed an all-pole model for the representation of acoustic signal frames. For an all-pole model LP parameters are the solution of the set of linear equations. All the simulation has been carried out in MATLAB. Autocorrelation method was used to get LP parameters. Acoustic signals from helicopter and non-helicopter sources were partitioned into training and test sets. The training set consists of 340 frames, including 200 helicopter sound frames and 140 non-helicopter sound frames. For a 0.1 second frame, different lengths (10,20,30) of LP coefficient features were tested. Among them a 20-coefficient feature vector has shown better convergence rate as well as a better performance. Of course, the convergence rate depends upon the initialization of the network weights, which in our study were chosen randomly.

Number of training epochs and hidden units in the ANN classifiers were varied in order to experimentally determine the optimal combination. Only one hidden layer

of neurons was used. A sigmoidal tangent non-linearity was used in hidden layer and a linear function was assumed in the output layer. Helicopter signals were assigned a target of 0.9 while non-helicopter signals were assigned a target of 0.1. In the testing phase, a threshold of 0.5 is used for deciding a helicopter or non-helicopter signal. The number of hidden neurons was varied between 5 and 15. An ANN was trained to a target sum squared error (SSE) goal of 0.1. Number of neurons in input layer were equal to number of features representing the acoustic signal while the output layer contained one neuron, equal to number of target bits. As the network size was varied by changing the number of neurons in the hidden layer, 10 hidden neurons have reached the minimum SSE goal of 1.3 when no further change in the SSE was observed.

The performance of the classifier on the test data is an indication of real identification performance. Although, the training and test data are drawn from the same class of populations, since the classifier is optimized on the training data, it might overlearn the training data missing more relevant discriminating characteristics. So classification error may increase on the test data if the class populations were undersampled or misrepresented in the training data.

The following figures (Fig.(7.1-7.4)) show the detection results using LP coefficients as features when the test data was corrupted with additive noise. These results are also shown in Table 7.1. Performance of the classifier is plotted against signal strength in Fig.(7.5).

Table 7.1: Correct detection using the LPC coefficients

Detection Results of	SNR				
	Clean	18 dB	12 dB	9 dB	3 dB
Helicopter	99.5 %	88 %	60 %	47.5 %	20.5 %
Non-helicopter	100 %	99 %	100 %	100 %	100 %
False Alarm	0 %	1 %	0 %	0 %	0 %
Miss	0.5 %	11 %	40 %	52.5 %	79.5 %

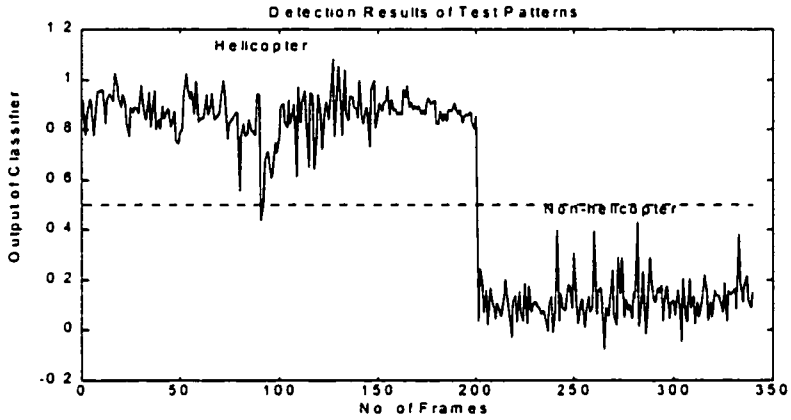


Fig.(7.1): LPC based ANN classifier performance with clean signal

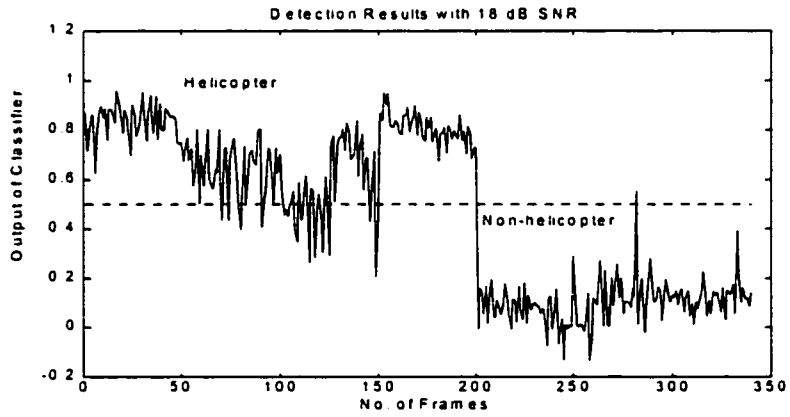


Fig.(7.2): LPC based ANN classifier performance with 18dB SNR

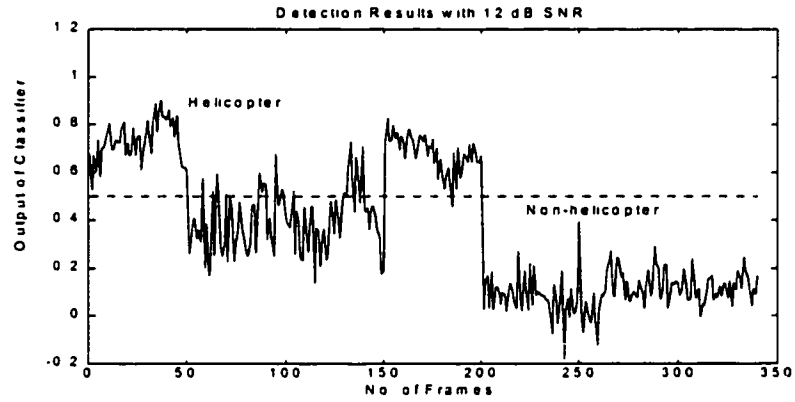


Fig.(7.3): LPC based ANN classifier performance with 12dB SNR

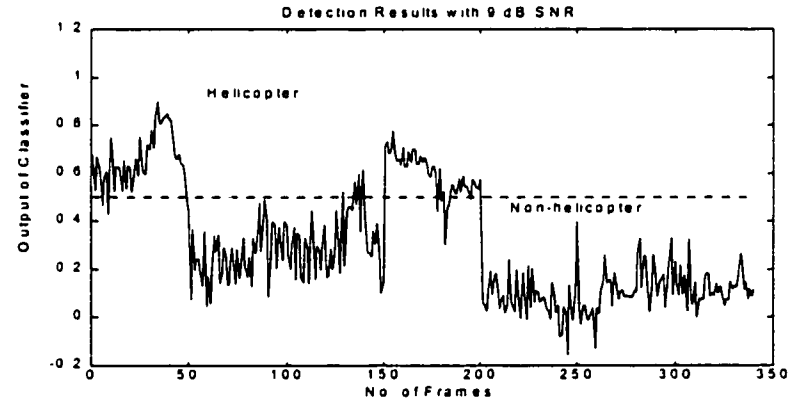


Fig.(7.4): LPC based ANN classifier performance with 9dB SNR

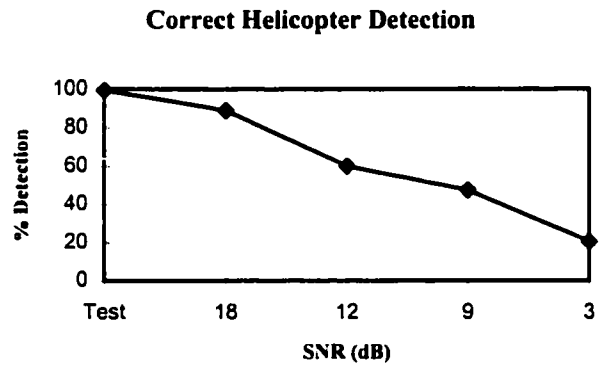


Fig.(7.5): LPC based detection with varying SNR

When the conditions were identical to those encountered in the training set, i.e. no noise is added to the test patterns, there was a small drop in recognition performance. This indicates that the training data has accounted for most of the variability in the test data. The performance of ANN classifier was also evaluated when the signal was embedded in the background noise. Data was generated from test frames corrupted by Gaussian noise with 18dB, 12dB and 9dB signal-to-noise ratio (SNR). The performance of the recognition system was affected as the spectrum shape distorts and becomes more buried in the background noise. As the signal strength drops, relative to the background noise, the spectral peaks blend in the noise that distorts the spectral shape, until it becomes very different from its original shape. In response the feature values change, affecting the performance quality. The degradation shown in the results illustrates the sensitivity of the ANN to these spectral changes.

7.3 Detection Using the Reflection Coefficients as Features

Reflection coefficients constitute another parameterization of the acoustic spectrum. Reflection coefficients have a much smaller sensitivity to finite word length (FWL) computation. In addition to the ease of computation, stability under quantization and natural ordering, the values of the reflection coefficients do not change as the filter length is increased. The stability of the filter is easily checked using the following condition

$$|k_i| < 1, \qquad 1 \le i \le p$$

These inherent properties of reflection coefficients have encouraged us to use them as feature vectors for a recognition system. Among different number of coefficients investigated, feature vectors with 30 reflection coefficients have shown faster convergence. Each 0.1 second frame was parameterized in terms of 30 reflection coefficients. Alternate frames were selected to make the training set of data. 540 frames, 400 helicopter and 140 non-helicopter frames were engaged in the training. Three different networks with one hidden layer containing 10, 15 and 20 neurons were trained to a SSE goal of 0.1. Due to small dynamic range $|k_i| < 1$, and natural ordering of the reflection coefficients, all the three networks have shown much faster convergence than the LP coefficients based scheme.

The classification ability of the networks was checked on test patterns without noise, and patterns with background noise of SNR 18dB, 12dB and 9dB. The results in terms of correct classifying ability and miss are shown in Table 7.2 below. It is clear that the performance has upgraded a bit in terms of correct detection as the size of the network is increased, but computational overhead and training time is increased. The correct detection is also plotted against the signal strength in Fig(7.6).

Here again it was noticed that the performance of the ANN classifier was affected by the increasing background noise. For the test patterns all the three classifiers (i.e. 10, 15 and 20 hidden neurons) have shown 100% correct detection when no noise was added. The degradation in performance in the presence of noise is due to the spectral sensitivity property of the reflection coefficients.

Table 7.2: Correct detection using the reflection coefficients

No. of Hidden layer neurons		Helicopter Detection	Non-Helicopter Detection	False Alarm	Miss
10	Clean	100 %	98.5 %	1.5 %	0 %
	18 dB	98.5 %	98.5 %	1.5 %	1.5 %
	12 dB	63 %	98.5 %	1.5 %	37 %
	9 dB	39.5 %	100 %	0 %	60.5 %
15	Clean	100 %	100 %	0 %	0 %
	18 dB	98.5 %	98.5 %	1.5 %	1.5 %
	12 dB	62.5 %	100 %	0 %	37.5 %
	9 dB	37.5 %	100 %	0 %	62.5 %
20	Clean	100 %	100 %	0 %	0 %
	18 dB	98.5 %	100 %	0 %	1.5 %
	12 dB	65.5 %	100 %	0 %	34.5 %
	9 dB	40 %	100 %	0 %	60 %

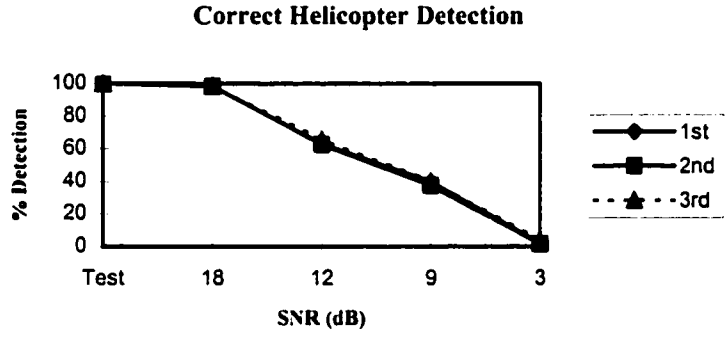


Fig.(7.6): Reflection coefficients based detection with varying SNR

7.4 Detection Using LSPs as Features

LSPs are another parameter set carrying kind of spectral information that resolves the effect of inappropriate quantization and filter instability problems of the LPC coefficients. LSP parameters have both well-behaved dynamic range and filter stability preservation property. They can be used to encode LPC spectral information

more efficiently than reflection coefficients and log-area ratios. LSP parameters are interpretable in terms of the formant frequencies of the model. The zeros of the LSP polynomial lie on the unit circle and the roots of the symmetric and anti-symmetric polynomial are interlaced. LSPs have shown some promise in the field of speaker recognition. This motivates to investigate its application as well in helicopter detection.

In our study, each 0.1 second acoustic frame was encoded into a 10 point LSP feature vector. 270 frames of training data, 200-helicopter sound frames and 70 non-helicopter sound frames were produced on keep-one-leave-three basis from the available acoustic signals. Due to the small number of LSP input features (10 in our case, because of the limitation of the root finding algorithm used) and the sensitivity of parameters to spectral variations, two hidden layers were employed to capture the spectral contents of the signal. Hidden layers contained 8 and 4 neurons respectively. The network was trained to a SSE of 1.3, after which no change in SSE was observed. The detection results on test data are found to be good but on noisy signal performance has gradually dropped, as the signal strength with respect to noise becomes low. The distribution of the differences between different LSP frequencies of the same utterance by different people generally lie within a small range (0-0.1 normalized frequency) for the 10^{th} order LSP analysis of speech frames [46]. This indicates that their statistical characteristics remain practically the same for slight deviation from the ideal conditions. However, sever alteration in conditions contribute considerable variations in the corresponding LSP frequencies resulting in

performance degradation. The helicopter detection results using LSPs are shown in Table 7.3 below and are also plotted against the signal strength in Fig.(7.7).

Table 7.3: Correct detection using LSPs

Detection Results of	SNR				
	Clean	18 dB	12 dB	9 dB	3 dB
Helicopter	100 %	99 %	83.5 %	69.5 %	34.5 %
Non-helicopter	98.5 %	100 %	100 %	100%	100 %
False Alarm	1.5 %	0 %	0 %	0 %	0 %
Miss	0 %	1 %	16.5 %	30.5 %	65.5 %

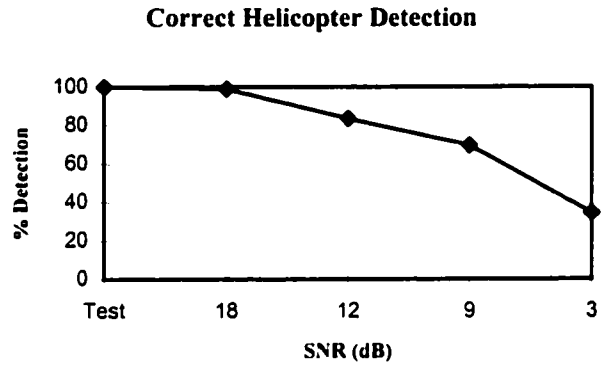


Fig.(7.7): LSP based detection with varying SNR

The results are indicative that LSPs have performed better in terms of correct classification. This is due to the fact that LSPs have a low spectral sensitivity as compared to direct LPC parameters. Decrease in detection performance is more due to the weakness of ANNs in dealing with noise than with LSPs spectral sensitivity.

7.5 Detection Using Cepstrum Coefficients

In the field of speaker recognition, cepstrum coefficients and their derivatives have been found to produce good results. They capture the temporal information essential

for text-dependent recognition task. In our application acoustic frames were represented in the form of 20 point cepstrum feature vector. An ANN with 12 neurons in one hidden layer was trained to a SSE goal of 0.1 using 270 training frames. Subsequent to training, the classifying ability of the recognizer was tested. For the clean data the detection accuracy has been 100%. However, the detection accuracy deteriorated with noise addition. The results are shown in Table 7.4 and are also plotted in Fig.(7.8).

Table 7.4: Correct detection using cepstrum coefficients

Detection Results of	SNR			
	Clean	18 dB	12 dB	9 dB
Helicopter	100 %	95.5 %	64 %	38 %
Non-helicopter	100 %	100 %	100 %	100 %
False Alarm	0 %	0 %	0 %	0 %
Miss	0 %	4.5 %	36 %	62 %

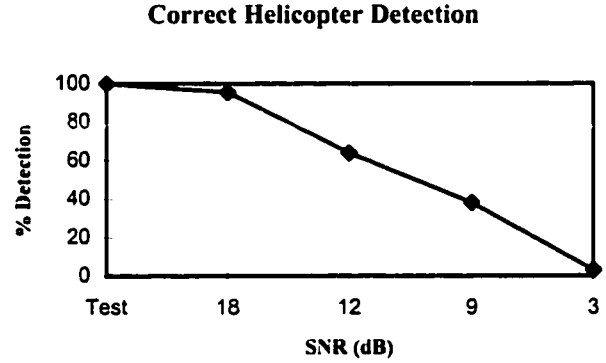


Fig.(7.8): Cepstrum coefficients based detection with varying SNR

The 'side effect' of non-linear log operation is the potential for improper emphasis of the low-level noisy portion of the acoustic spectrum. This fact has degraded the performance of recognition strategy based on cepstrum coefficients[78].

7.6 Comparison of the Feature Extraction Techniques

In this section a composite performance comparison of different feature extraction techniques in ANN based helicopter detection schemes is presented. Detection accuracy of four feature extraction techniques described earlier for 12dB and 9dB SNR is shown in Tables 7.5 and 7.6 below.

Table 7.5: Detection results using different schemes at 12dB SNR

Detection Results of	Encoding Scheme Used			
	LPC Coeffs.	Ref. Coeffs.	LSPs	Cepstrum Coeffs.
Helicopter	60 %	65.5 %	83.5 %	64 %
Non-helicopter	100 %	100 %	100%	100 %
False Alarm	0 %	0 %	0 %	0 %
Miss	40 %	34.5 %	16.5 %	36 %

Table 7.6: Detection results using different schemes at 9dB SNR

Detection Results of	Encoding Scheme Used			
	LPC Coeffs.	Ref. Coeffs.	LSPs	Cepstrum Coeffs.
Helicopter	47.5 %	40 %	69.5 %	38 %
Non-helicopter	100 %	100 %	100%	100 %
False Alarm	0 %	0 %	0 %	0 %
Miss	52.5 %	60 %	30.5 %	62 %

In Fig.(7.9), the detection ability of the best classifier of each coding technique is plotted against the SNR. The results are indicative that LSP's performed best in terms of correct classification.

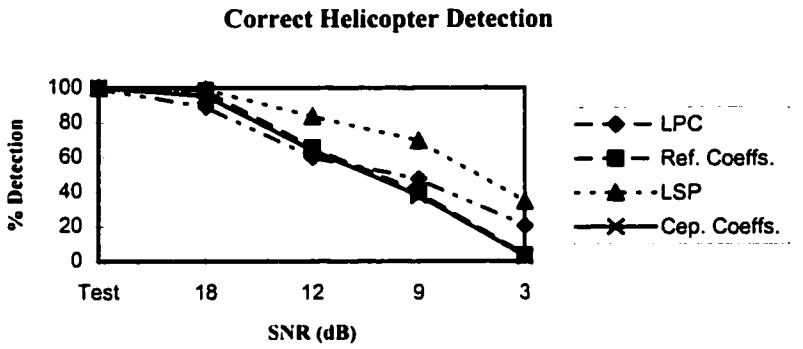


Fig.(7.9): Detection performance of four schemes with varying signal strength

Up till now we have applied four different techniques for obtaining acoustic feature vectors. It has been observed that addition of background noise in the test signal has degraded the detection performance. To counteract the effect of noise two approaches were investigated:

- De-noising of the signal prior to feature extraction
- Training the system in an environment similar to test the conditions

7.7 De-Noising Of The Signal Prior To Feature Extraction

Due to substantial research in the signal processing area in last two decades, a new technique called wavelet analysis has emerged. This technique was applied for de-noising of acoustic signal prior to finding features as shown in Fig.(7.10). The wavelet analysis is a distinct improvement on the Short time Fourier Transform (STFT). It represents a signal in terms of time and frequency.

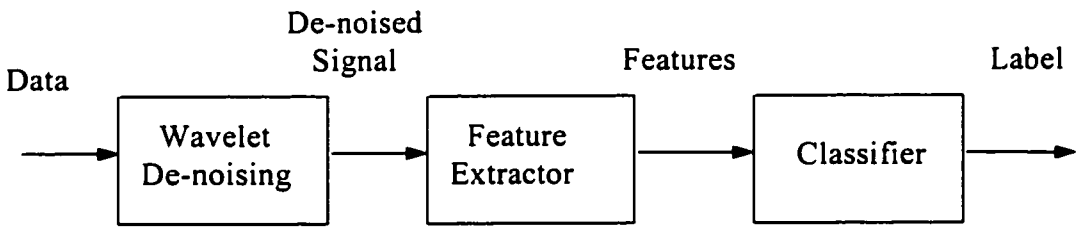


Fig.(7.10): Block diagram showing de-noising prior to feature extraction

Depending upon the problem a continuous or discrete wavelet transform can be applied. At the first level wavelet transform partitioned the data into two half bands, a low frequency band and a high frequency band. These bands can be further partitioned into two half bands and so on. In this way the signal can be resolved at various resolutions. The high frequency components are assumed to carry noise, so they can be eliminated following some criteria defining the threshold. The Figs.(7.11-7.13) below show de-noised signals using wavelets.

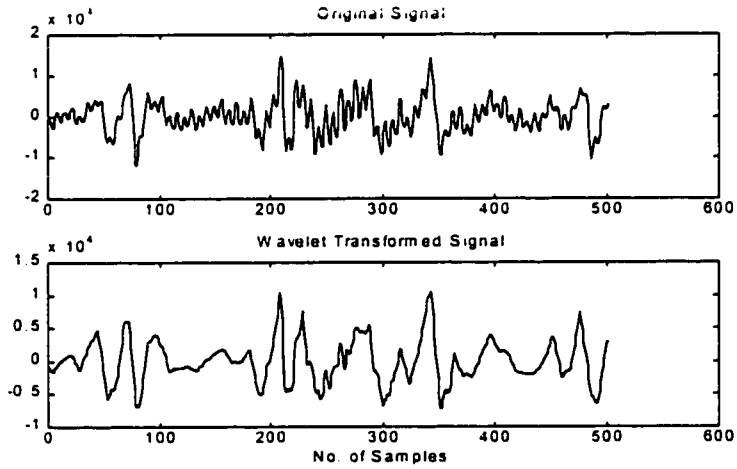


Fig.(7.11): Clean signal and its de-noised version

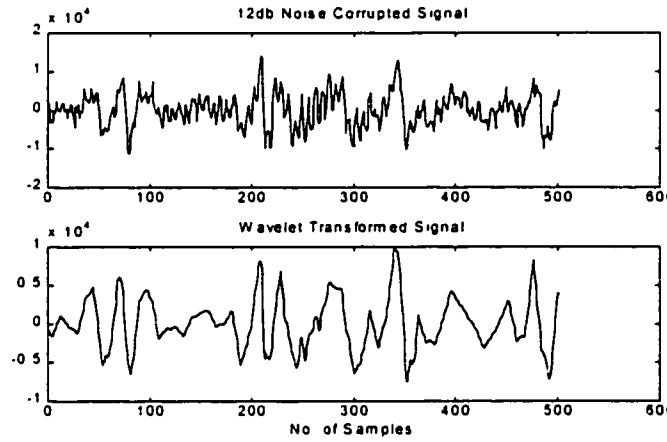


Fig.(7.12): 12dB noise corrupted signal and its de-noised version

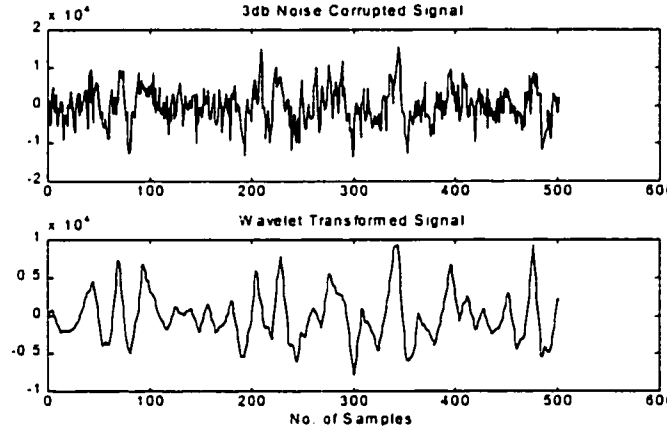


Fig.(7.13): 3dB noise corrupted signal and its de-noised version

In our application we have used discrete wavelet transform. MATLAB Wavelet Toolbox functions "DDENCMP and WDENCMP" were used for threshold selection and de-noising respectively. Compactly supported orthonormal wavelet "db3" was used in the analysis. A three level decomposition of the signal is performed as shown in Fig.(7.14). By successive approximation the signal becomes less and less noisy, however it may as well lose some high frequency component of the signal. For noise

removal we have to discard some of the high frequency information. An entropy-based scheme to select the global threshold is applied.

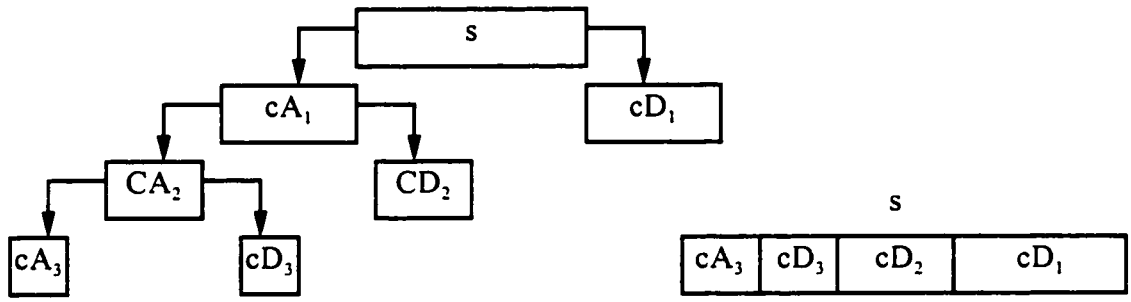


Fig.(7.14):Block diagram showing 3 level decomposition of signal using wavelets

From this de-noised signal spectral estimates were obtained in the form of 20 reflection coefficients. ANN classifier was trained using these feature vectors. Two different classifiers with two hidden layers of neurons, first having 10 and 5 neurons and second having 18 and 12 neurons, were trained. The results obtained are shown in the Table 7.7 below. These results illustrate that there is an improvement in the performance compared to the earlier results. The detection performance with varying signal strength is also plotted in Fig.(7.15).

Table 7.7: Correct detection using de-noised signal

No. of hidden layer neurons		Helicopter Detection	Non-Helicopter Detection	False Alarm	Miss
10 and 5	Clean	100 %	100 %	0 %	0 %
	18 dB	88.5 %	92.7 %	7.3 %	11.5 %
	12 dB	73 %	86.5 %	13.5 %	27 %
	9 dB	59.5 %	92 %	8 %	40.5 %
18 and 12	Clean	100 %	100 %	0 %	0 %
	18 dB	90 %	92 %	8 %	10 %
	12 dB	75.5 %	86.4 %	13.6 %	24.5 %
	9 dB	60 %	92 %	8 %	40 %

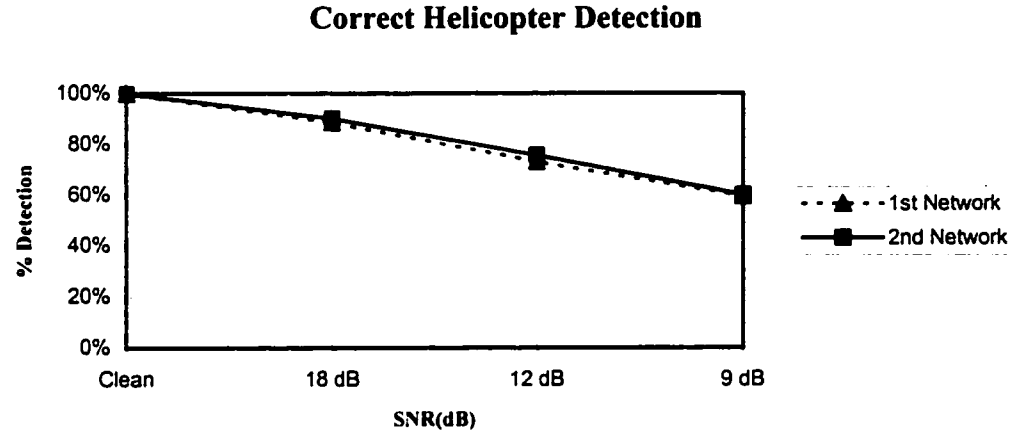


Fig.(7.15): Classifier performance using de-noised signal

7.8 Training The System In Environment Similar To Test Conditions

Another way to make the ANN less sensitive to changes in spectral shapes due to additive background noise is to train the network using templates created with background noise. This conjecture has been tested through simulation study. In this simulation an ANN classifier was trained on 810 frames, 540 frames without noise and 270 frames with additive noise level of 9dB SNR. The ANN had one hidden layer of neurons having 13 neurons. Feature vector is obtained in the form of 30 reflection coefficients. Later on, the performance of the classifier was tested on test frames (without noise) and 18dB, 12dB, 9dB and 3dB SNR noise corrupted frames not included in the training. The results in Table 7.8 shows noticeable improved performance even in case of heavy background noise. The detection results are shown in Figs.(7.16-7.19) and are also plotted for different SNR in Fig.(7.20).

Table 7.8: Correct detection using clean and noisy data for training

Detection Results of	Signal Applied to the Classifier				
	Clean	18 dB	12 dB	9 dB	3 dB
Helicopter	100 %	100 %	100 %	100 %	88 %
Non-helicopter	98.5 %	97 %	100 %	100 %	100 %
False Alarm	1.5 %	3 %	0 %	0 %	0 %
Miss	0 %	0 %	0 %	0 %	12 %

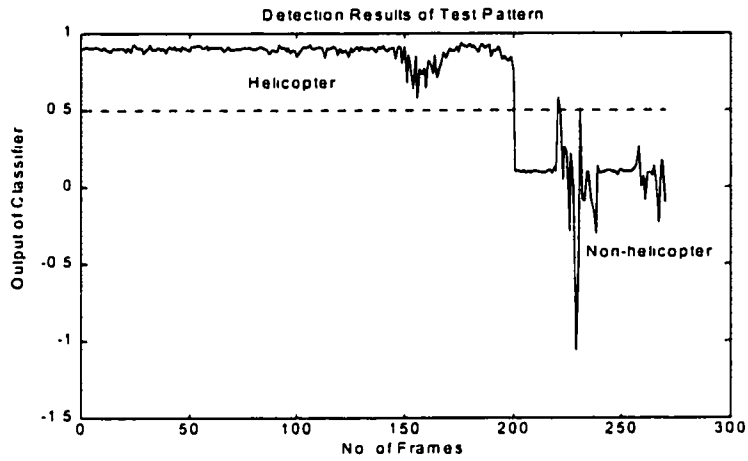


Fig.(7.16): ANN classifier performance with clean signal

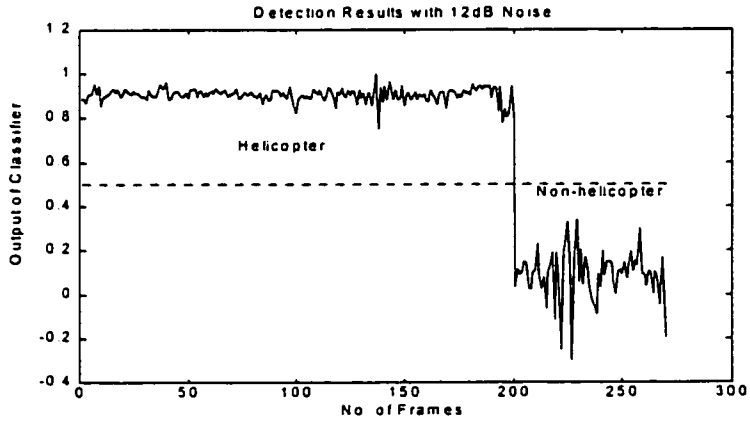


Fig.(7.17): ANN classifier performance with 12dB SNR

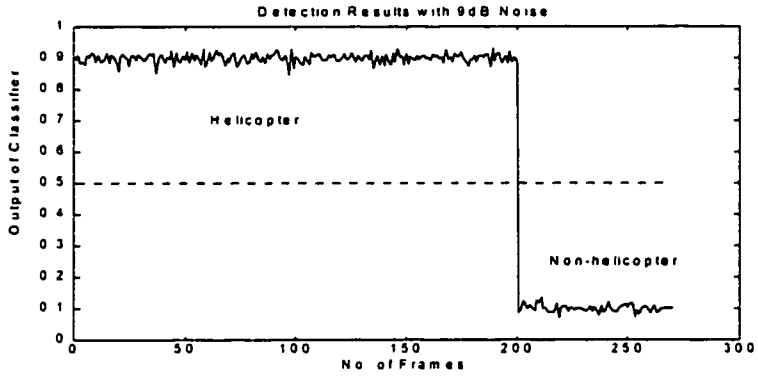


Fig.(7.18): ANN classifier performance with 9dB SNR

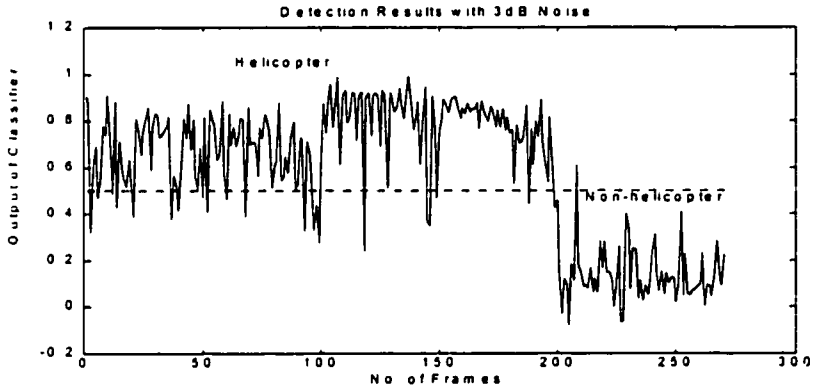


Fig.(7.19): ANN classifier performance with 3dB SNR

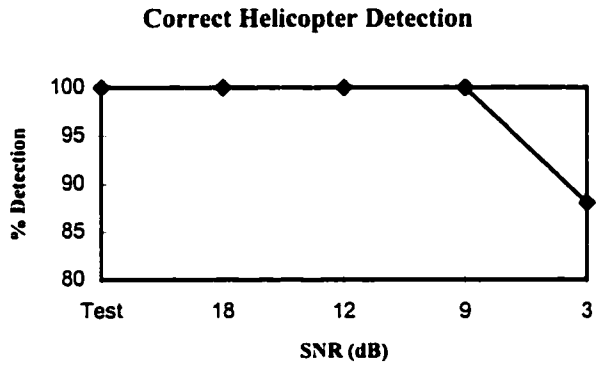


Fig.(7.20): ANN classifier performance with varying SNR

The plot below compares the detection performance of the best coding technique (LSP) trained on clean signal only, best wavelet de-noised classification and the performance of the classifier which was trained on noisy as well as clean signal. It is clear that the performance of the latest is far better than the others.

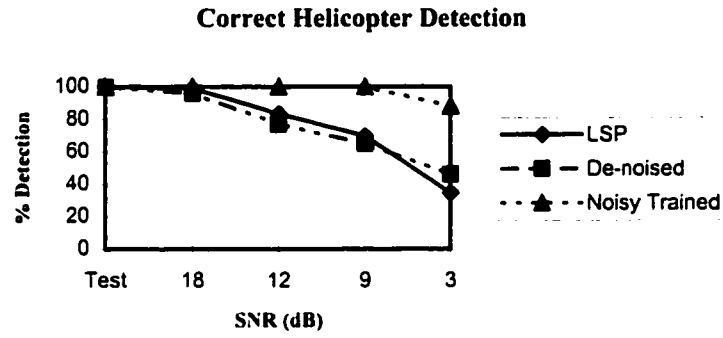


Fig.(7.21): Performance comparison of three ANN classifiers

We have noticed that the performance of the LPC and its derivative based classifiers are degraded due to background noise. In case of LPC features, the noise tends to reduce the dynamic range of the resulting coefficients. The peaks and valleys of the spectrum become unclear due to noise [79]. For log spectrum (cepstrum coefficients) the noise has larger effect in the spectral valleys than the spectral peaks which deteriorates the detection performance. Due to relatively small dynamic range of the reflection coefficients, additive noise of relatively low power produce small variations in the reflection coefficients resulting in satisfactory performance. They have also been used in the subsequent performance improvement studies because of the fast convergence rate of the ANN. The statistical analysis of LSPs has shown slight variation in formant frequencies with changing SNR [46]. This results in better performance than other encoding techniques used in this study.

For LPC and reflection coefficients we have used high dimensional input features, as ANN classifier with lower number of input features did not converge. For LSP and cepstrum coefficient feature vectors, networks were taking too long to converge (they were taking more than two weeks as compared to a few days for LPC and reflection coefficient input features on Pentium 200 computers. Moreover, in attempts to get higher order input features for the LSP, the root finding algorithm became unstable. To make the use of LSP features practicable faster computers as well as better root finding algorithm are required.

Chapter 8

Helicopter Classification: Simulation Studies

Summary

This chapter describes the results of the simulation study on the classification of helicopters using acoustic information. ANN pattern classifiers were trained using different feature extraction techniques. The performance of these classifiers on test data is reported. A classifier was also trained using different flight conditions. Wavelet de-noising capability is applied to enhance the performance in noisy signal conditions. The performance of a classifier trained on clean and noisy signal to differentiate among eight different helicopters is also presented.

8.1 An Hierarchical Detection and Classification System

Hierarchy of neural network classifiers is used in this detection and classification problem. At the first level an ANN recognition system decides about the presence of helicopter. Once a helicopter is detected, a second network differentiates between the helicopter types. In this study we have used simulated sound of four helicopters S-67, H500, CH-47C and MI-HIND abbreviated as A, B, C, D respectively. Each of these helicopters has a different characteristic sound due to different rotor blade diameters, number of rotor blades, chord length, engine power and disc loading factor etc. The motion of the helicopter causes continuous change in the shape of the spectrum, some times the spectral estimates of different helicopter sounds are very close to each other. This additive complexity has made the classification job more arduous.

According to the "Universal Approximation Theorem", a single hidden layer is sufficient for a multilayer perceptron to compute a uniform approximation to a given training set. There is no restriction on the number of neurons in this single layer. The problem of using a single hidden layer is that the neurons tend to interact with each other. In complex situations this interaction makes it difficult to improve the approximation at one point without worsening it at some other point. By the use of two hidden layers, the approximation process becomes more manageable. Local features are extracted in first hidden layer. Some neurons in the first hidden layer are used to partition the space into two regions. Other neurons in this layer learn the local features characterizing that region. Global features are extracted in the second

hidden layer [73]. Preliminary studies regarding this classification problem has revealed that single hidden layer is unable to handle underlying complexity of the problem. Therefore two hidden layers were used in the classifier design.

Spectral estimates are obtained in the form of LPC coefficients, reflection coefficients, LSPs, and cepstrum coefficients. Two output neurons were used and helicopters A, B, C, D are assigned targets of (0.1 0.1), (0.1 0.9), (0.9 0.1), (0.9 0.9) respectively. The classification performance of each feature extraction technique is discussed in the following sections.

8.2 Classification Using the LPC Coefficients

As a preliminary study, the acoustic data for four helicopters is obtained considering four different flight conditions. These different flight conditions depend upon helicopter height, helicopter velocity and its x-axis and y-axis distances form the observer.

400 frames were included in the training, 25 for each flight condition for each helicopter. Feature vectors are obtained in the form of 30 LPC coefficients. Training frames were chosen leaving 7 frames between the consecutive selections. Three different networks were trained to a SSE goal of 0.1. The performance of these classifiers on test and noisy data are shown in Table 8.1 and are also shown in Figs.(8.1-8.3).

It can be observed that the performance of the classifier was degraded as the noise level increased. With the increase in noise the spectral peaks and valleys

become unclear and sometimes disappear. The noise tends to reduce the dynamic range and variance between the frames, which results in performance degradation.

Table 8.1:Classification results using the LPC coefficients

No. of hidden layer neurons		Correct Classification of			
		Helicopter A	Helicopter B	Helicopter C	Helicopter D
12 & 5	Clean	100 %	100 %	100 %	100 %
	12 dB	97 %	86 %	90 %	90 %
	9 dB	88 %	59 %	79 %	92 %
	3 dB	57 %	25 %	58 %	75 %
15 & 8	Clean	100 %	100 %	100 %	100 %
	12 dB	97 %	79 %	93 %	92 %
	9 dB	90 %	49 %	83 %	92 %
	3 dB	65 %	10 %	68 %	75 %
17 & 10	Clean	100 %	100 %	100 %	100 %
	12 dB	99 %	99 %	66 %	93 %
	9 dB	96 %	95 %	48 %	90 %
	3 dB	84 %	59 %	14 %	77 %

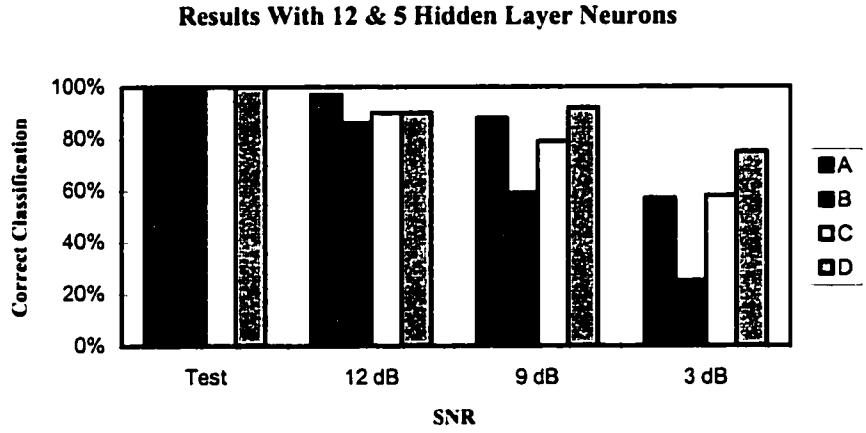


Fig.(8.1): Performance of 1st LPC based classifier

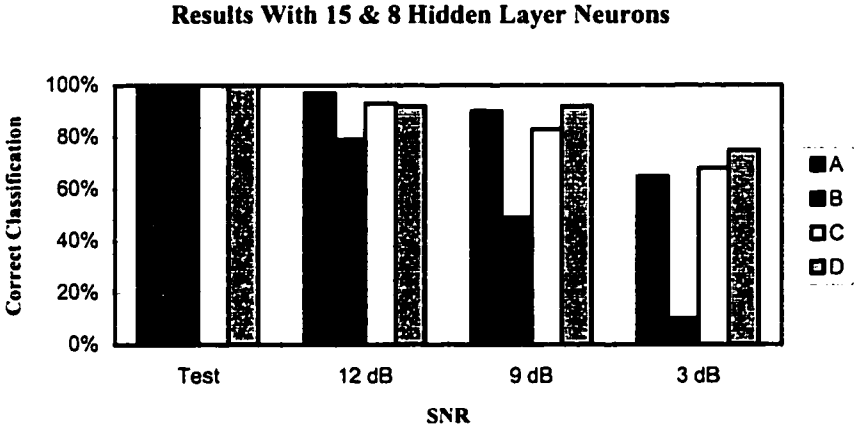


Fig.(8.2): Performance of 2nd LPC based classifier

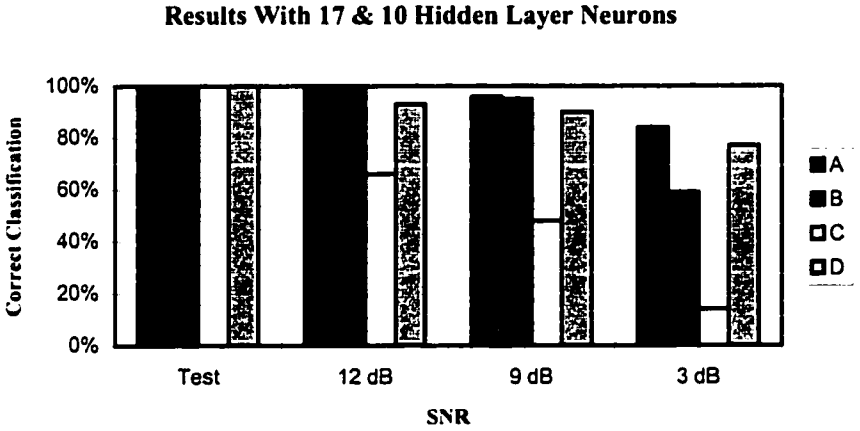


Fig.(8.3): Performance of 3rd LPC based classifier

8.3 Classification Using the Reflection Coefficients as Features

Four different networks were trained using the reflection coefficients as feature extractor. Each 0.1 second helicopter acoustic signal frame was encoded in 30 reflection coefficients. Training data consists of 400 frames, 100 from each helicopter. Training frames were selected on keep-one-leave-one basis. The networks

were trained to a SSE goal of 0.1. The performance of the trained networks are evaluated using test frames not included in the training. Classifiers were also tested when the training frames were corrupted by background noise of 12dB, 9dB, 3dB SNR. The classification performance is summarized in Table 8.2.

The results have shown that the network having 15 and 7 neurons has produced the most acceptable results in this case. The results are also plotted in Figs.(8.4-8.7) below. The classifying ability of each network is also given in Figs.(8.8-8.11) relative to the signal strength. The performance is gradually decreased when the network size is increased. This is due to the fact that added neurons have captured the characteristics which are not common to the entire populations. Thus determining an optimal network size is a matter of trial and error.

Table 8.2: Classification results using the reflection coefficients

No. of hidden layer neurons		Correct Classification of			
		Helicopter A	Helicopter B	Helicopter C	Helicopter D
20 & 10	Clean	100 %	100 %	100 %	100 %
	12 dB	100 %	100 %	99 %	95 %
	9 dB	100 %	100 %	96 %	92 %
	3 dB	100 %	87 %	94 %	33 %
15 & 7	Clean	100 %	100 %	100 %	100 %
	12 dB	100 %	100 %	100 %	100 %
	9 dB	100 %	100 %	100 %	100 %
	3 dB	100 %	91 %	100 %	100 %
7 & 3	Clean	100 %	100 %	100 %	100 %
	12 dB	100 %	100 %	100 %	70 %
	9 dB	100 %	100 %	100 %	52 %
	3 dB	100 %	99 %	100 %	43 %
5 & 2	Clean	100 %	100 %	100 %	100 %
	12 dB	100 %	100 %	100 %	100 %
	9 dB	100 %	100 %	100 %	100 %
	3 dB	97 %	93 %	100 %	98 %

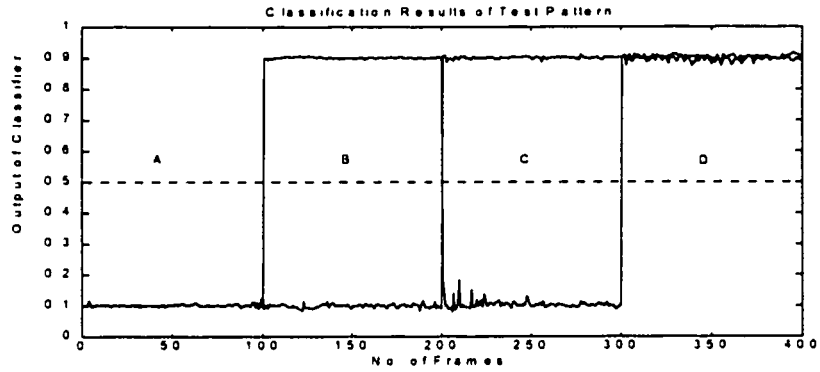


Fig.(8.4): Classifier performance with clean signal

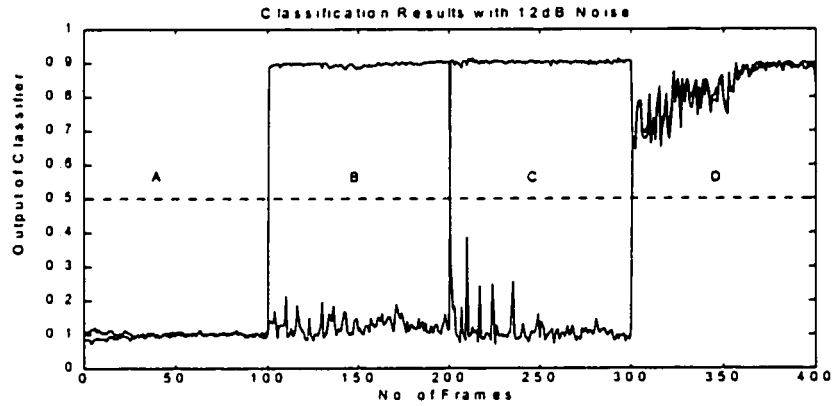


Fig.(8.5): Classifier performance with 12dB SNR

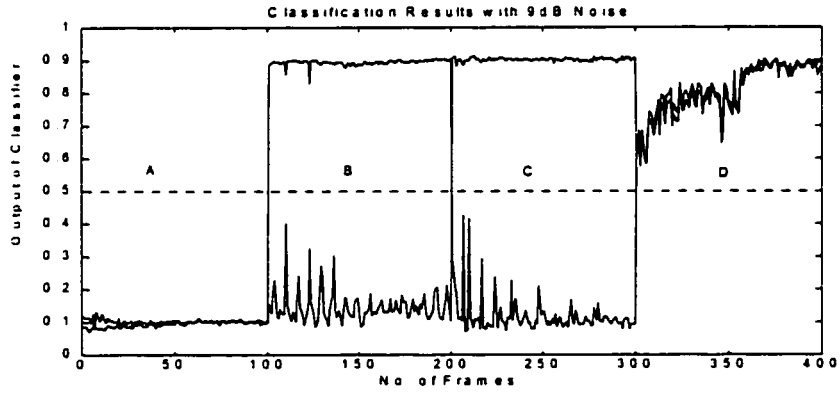


Fig.(8.6): Classifier performance with 9dB SNR

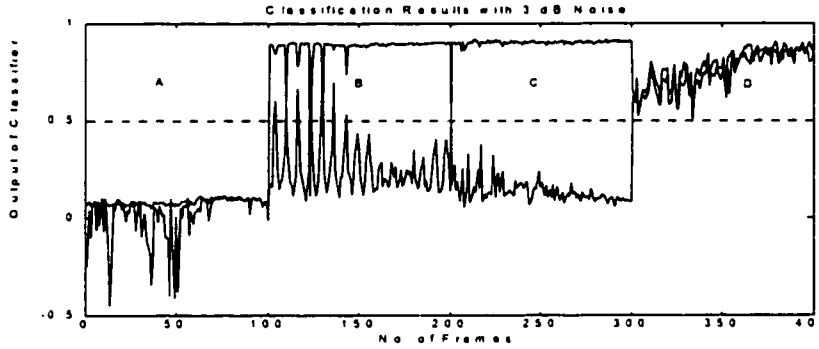


Fig.(8.7): Classifier performance with 3dB SNR

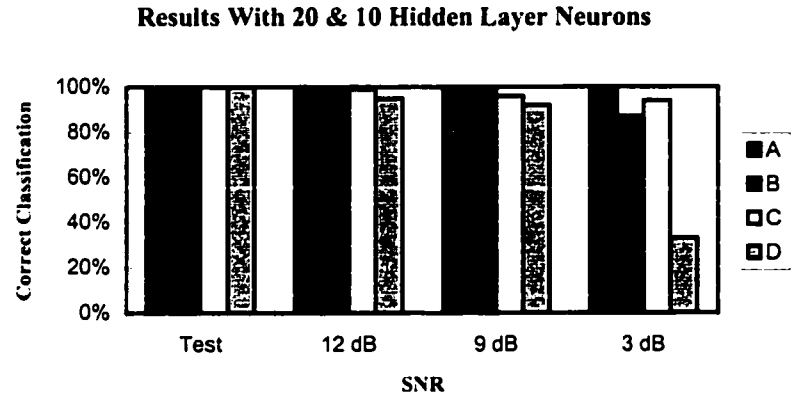


Fig.(8.8): Performance of 1st reflection coefficients based classifier

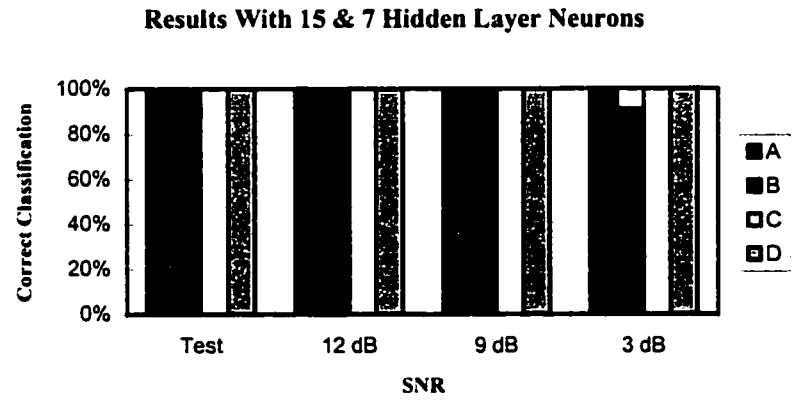


Fig.(8.9): Performance of 2nd reflection coefficients based classifier

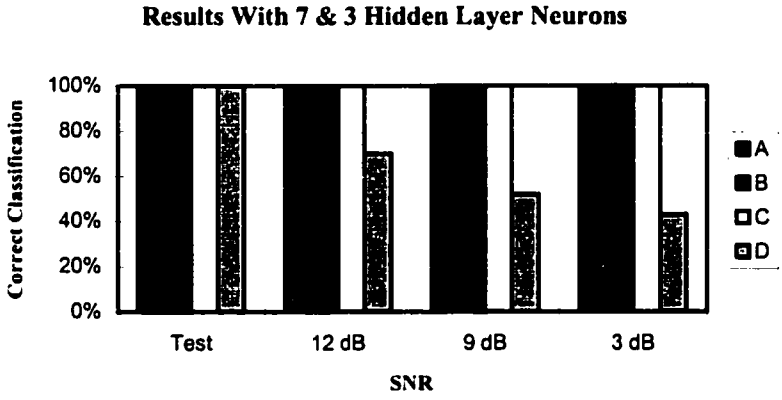


Fig.(8.10): Performance of 3rd reflection coefficients based classifier

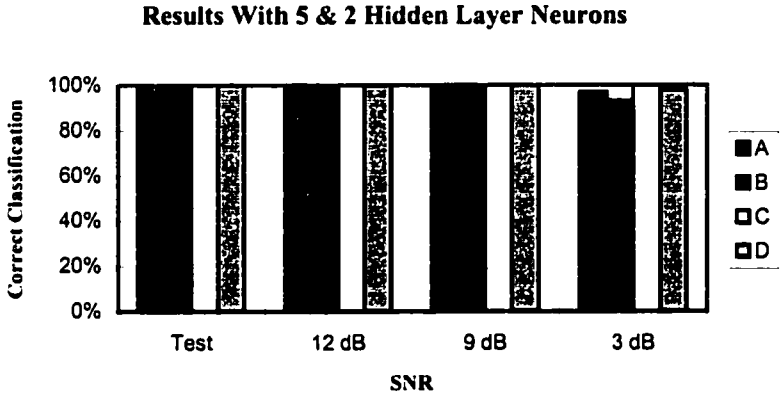


Fig.(8.11): Performance of 4th reflection coefficients based classifier

8.4 Classification Using LSPs as Features

In this simulation, ANN classifier was trained using LSPs as feature encoding technique. Each 0.1 second acoustic frame was translated into 10 LSP frequencies. Training set consists of 400 frames. ANN classifier had two hidden layers with 8 and 4 neurons in 1st and 2nd layer respectively. Training was done until a SSE goal of 0.1 is reached. The test results of the classifier on clean and noisy test frames are shown

in Table 8.3. With the increase in noise level, the performance of the classifier has been degraded. Due to the change in spectral shape with noise addition, the features drawn did not correlate to the one included in the training. The use of low order parameterization of the acoustic signal has also created some degradation in performance.

Table 8.3: Classification results using LSPs

Classification Results of Helicopter	SNR			
	Clean	12 dB	9 dB	3 dB
A	100 %	100 %	100 %	100 %
B	100 %	90 %	72 %	28 %
C	100 %	32 %	19 %	1 %
D	100 %	100 %	100 %	100 %

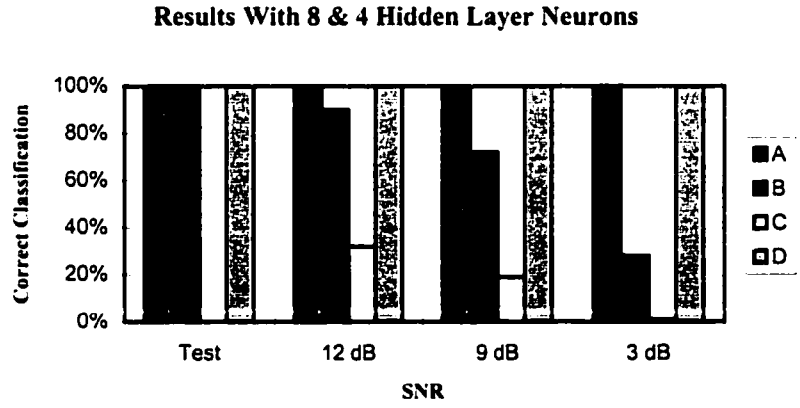


Fig.(8.12): Classification performance using LSP features

8.5 Classification Using Cepstrum Coefficients as Features

In this study, the ANN classifier having 6 and 3 neurons in two hidden layers was trained to a SSE goal of 0.1. Feature vectors are extracted from 0.1 second acoustic frame in the form of 10 cepstrum coefficients. Cepstrum coefficients are derived

from the reflection coefficients. The results of classifier test are shown in Table 8.4. Here again the results have shown that, increasing background noise has affected the performance, especially for helicopter A and B. It has been observed by other researchers [78] that noise did not effect the location of spectral peaks. Dynamic features are usually less effected. The difference between the true spectral model and the noisy spectral model is largely in the spectral valleys than in the spectral peaks. Lower order coefficients are more affected than the higher order terms [78].

Table 8.4: Classification results using cepstrum coefficients

Classification Results of Helicopter	SNR			
	Clean	12 dB	9 dB	3 dB
A	100 %	12 %	0 %	0 %
B	100 %	100 %	98 %	48 %
C	100 %	100 %	100 %	100 %
D	100 %	100 %	98 %	16 %

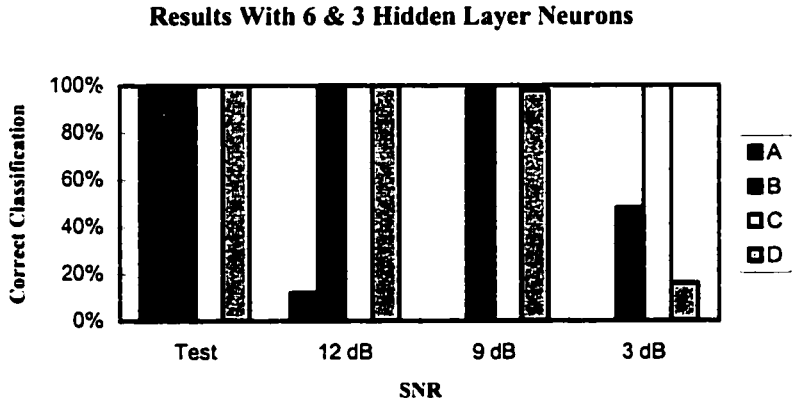


Fig.(8.13): Classification performance using cepstrum features

8.6 Comparison of Four Encoding Schemes

In the following tables the results of the above four encoding schemes are compared at 9dB and 3dB SNR. It is evident that among the four investigated encoding schemes; reflection coefficients have shown the best performance.

Table 8.5: Comparison of encoding techniques at 9dB

Classification Results of Helicopter	Encoding Scheme Used			
	LPC Coeffs.	Reflection Coeffs.	LSPs	Cepstrum Coeffs.
A	96 %	100 %	100 %	0 %
B	95 %	100 %	72 %	98 %
C	48 %	100 %	19 %	100 %
D	90 %	100 %	100 %	98 %

Table 8.6: Comparison of encoding techniques at 3dB

Classification Results of Helicopter	Encoding Scheme Used			
	LPC Coeffs.	Reflection Coeffs.	LSPs	Cepstrum Coeffs.
A	89 %	100 %	100 %	0 %
B	59 %	91 %	28 %	48 %
C	14 %	100 %	1 %	100 %
D	77 %	100 %	100 %	16 %

8.7 Use Of Wavelet Transform For Signal De-Noising

One way to reduce the detrimental effect of noise on the classification is to de-noise the signal during pre-processing. Then the spectral estimates were obtained from this de-noised signal. We have applied wavelet transform for de-noising the signal. Some

of the high frequency information in the signal was discarded by selecting the global threshold for de-noising. Discrete Daubechies wavelet "db3" was used in the analysis. Three levels of decomposition were used to encode the acoustic signal in approximation and details. Then this de-noised signal has been represented in the form of 20 reflection coefficients. 100 training frames of each helicopter constitute the 400 frames of training set. Four different networks with two hidden layers were trained to a SSE goal of 0.1. The performance of the classifiers were tested on clean signal, not included in the training, as well as on acoustic data corrupted by 12dB, 9dB and 3dB SNR.

The results in Table 8.7 have shown that the performance is improved with the increase in the network size up to a certain limit. With the increase in hidden neurons more relevant attributes of the training patterns are captured. However if there are too many hidden units, the increased non-linearity may introduce additional solutions in the weight space, complicating the learning procedure. Consequently, the performance may be degraded as we have encountered with a network with 23 and 12 hidden units.

The ANN classifier having 16 and 8 neurons in 1st and 2nd hidden layers respectively has produced the best results. The results were even better when we have used 30 reflection coefficients without de-noising the acoustic signal of each helicopter.

Table 8.7: Classification results applying wavelet de-noising capability

No. of hidden layer neurons		Correct Classification of			
		Helicopter A	Helicopter B	Helicopter C	Helicopter D
10 & 5	Clean	100 %	99 %	100 %	100 %
	12 dB	100 %	98 %	98 %	95 %
	9 dB	99 %	95 %	99 %	98 %
	3 dB	99 %	95 %	99 %	98 %
13 & 6	Clean	100 %	97 %	100 %	100 %
	12 dB	100 %	97 %	99 %	100 %
	9 dB	99 %	96 %	99 %	99 %
	3 dB	99 %	96 %	99 %	99 %
23 & 12	Clean	100 %	100 %	100 %	100 %
	12 dB	99 %	100 %	99 %	99 %
	9 dB	100 %	100 %	99 %	100 %
	3 dB	100 %	100 %	99 %	100 %
23 & 12	Clean	100 %	100 %	99 %	100 %
	12 dB	100 %	100 %	98 %	84 %
	9 dB	100 %	100 %	100 %	85 %
	3 dB	100 %	100 %	100 %	85 %

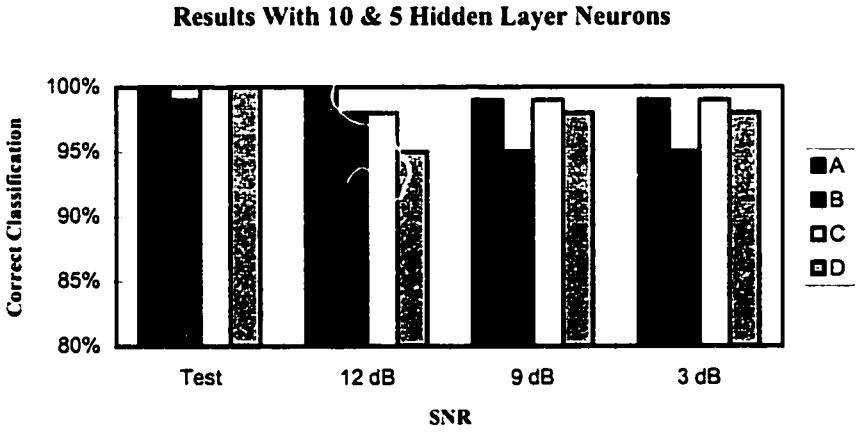


Fig.(8.14): Performance of 1st wavelet's de-noised classifier

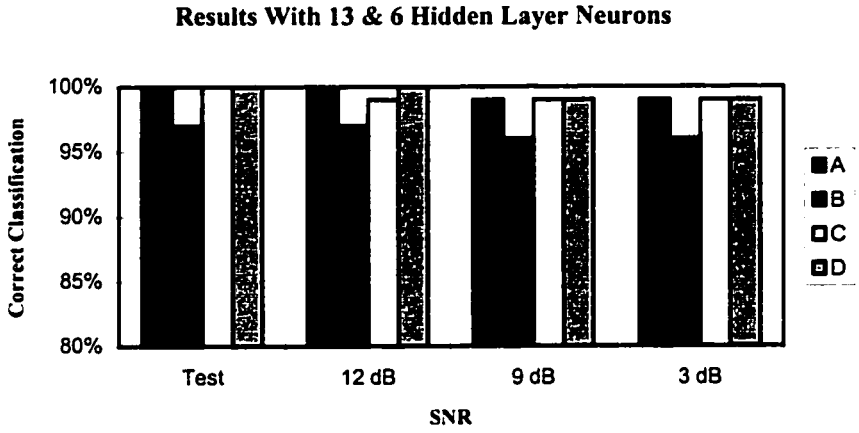


Fig.(8.15): Performance of 2nd wavelet's de-noised classifier

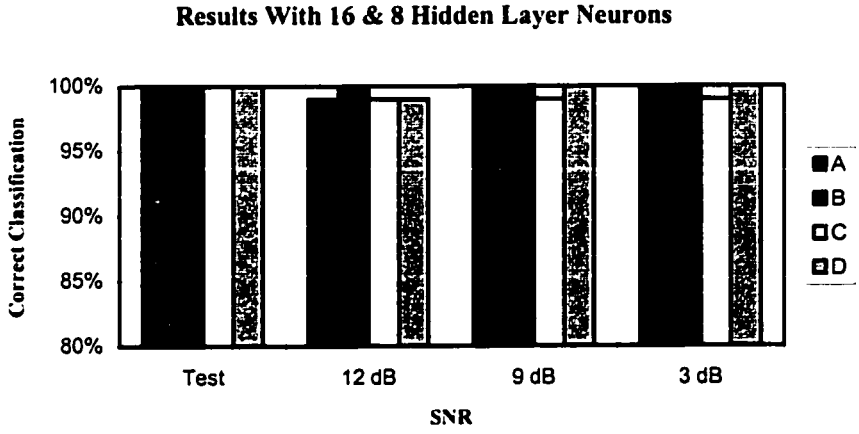


Fig.(8.16): Performance of 3rd wavelet's de-noised classifier

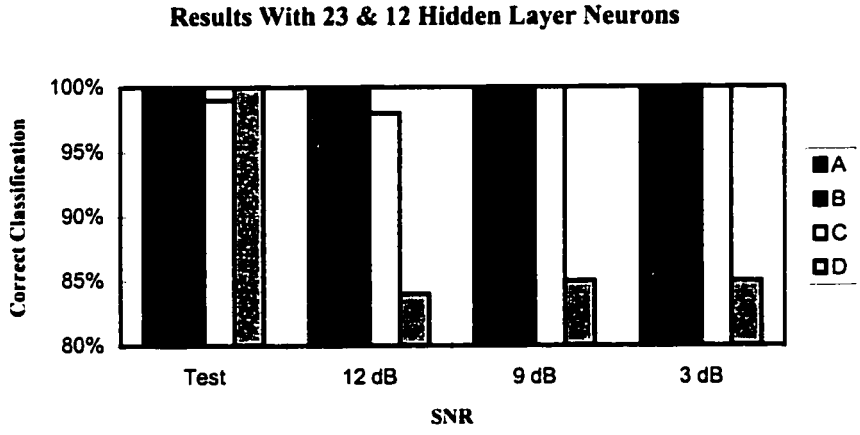


Fig.(8.17): Performance of 4th wavelet's de-noised classifier

8.8 Classification Of Eight Helicopters

A classifier was also trained to distinguish between 8 helicopters. The acoustic data from S-67, H500, CH-47C, HH-3, Mi-10, Mi-6, CH53 and MI-HIND helicopters were generated using the simulator described in chapter 3. Two different networks were trained. In first case, training data consists of 800 frames (100 for each helicopter) of clean signal only. For the second classifier, training set of data consisted of 800 frames. Among them, 400 frames consisted of clean signal and 400 frames were corrupted by 9dB noise. In each of the 400 frames, 50 frames of each helicopter were included. Frames were selected on keep-one-leave-seven terms. Helicopters A, B, C, ..., H respectively, has been assigned the columns of the following matrix as the ANN target values.

$$\begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.9 & 0.9 & 0.9 & 0.9 \\ 0.1 & 0.1 & 0.9 & 0.9 & 0.1 & 0.1 & 0.9 & 0.9 \\ 0.1 & 0.9 & 0.1 & 0.9 & 0.1 & 0.9 & 0.1 & 0.9 \end{bmatrix}$$

i.e. we had three neurons in the output layer, one for each target digit. Each acoustic frame was represented by 20 reflection coefficients. The classifiers had 10 and 5 neurons in the hidden layers. They were trained to a SSE goal of 0.1. Their performance was tested using 400 frames. The results on the clean and noisy test data are listed in Tables 8.8 and 8.9. The results show that the performance of the classifier is improved by addition of noisy frames in the training set as compared to the classifier trained on clean signal only.

Table 8.8: Classification results using clean signal in training

Classification Results of Helicopter	SNR					
	Clean	15 dB	12 dB	9 dB	6 dB	3 dB
A	100 %	100 %	100 %	100 %	96 %	96 %
B	100 %	48 %	22 %	0 %	4 %	0 %
C	96 %	100 %	100 %	100 %	98 %	94 %
D	98 %	82 %	84 %	74 %	70 %	74 %
E	98 %	26 %	12 %	4 %	0 %	0 %
F	98 %	82 %	78 %	72 %	60 %	50 %
G	100 %	2 %	2 %	0 %	2 %	0 %
H	100 %	98 %	98 %	92 %	80 %	66 %

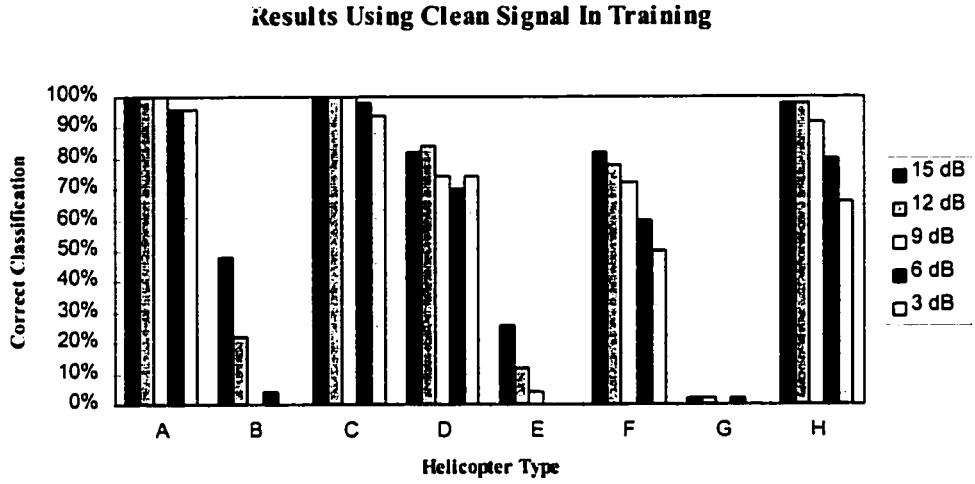


Fig.(8.18): Performance of the classifier trained on clean signal only

Table 8.9: Classification results using clean and noisy signals in training

Classification Results of Helicopter	SNR					
	Clean	15 dB	12 dB	9 dB	6 dB	3 dB
A	100 %	100 %	98 %	100 %	98 %	98 %
B	100 %	100 %	100 %	100 %	98 %	96 %
C	98 %	100 %	100 %	98 %	100 %	100 %
D	100 %	98 %	98 %	98 %	96 %	88 %
E	98 %	98 %	98 %	94 %	78 %	78 %
F	100 %	98 %	100 %	100 %	96 %	88 %
G	100 %	88 %	92 %	84 %	82 %	48 %
H	100 %	100 %	100 %	100 %	100 %	84 %

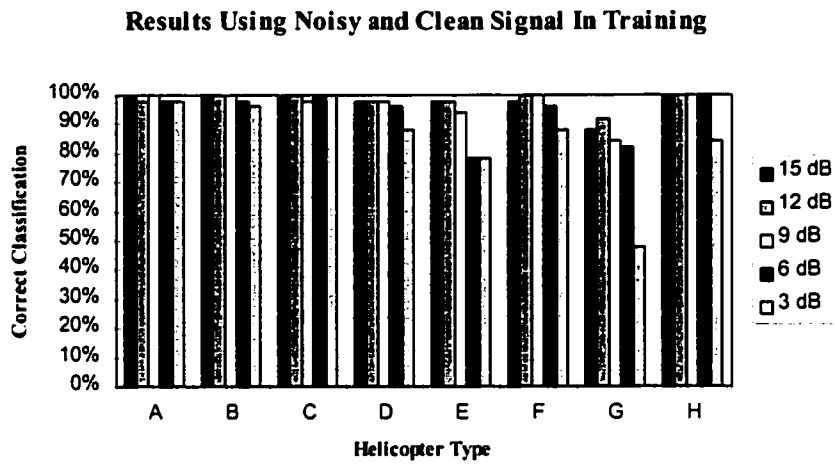


Fig.(8.19): Performance of the classifier trained on clean as well as noisy signal

When we have used wavelet transform for de-noising the signal at the pre-processing stage, we got better results. However, in this approach one has to put greater computational effort, and hardware cost to build the system would also increase. By the inclusion of noisy data in the training, the performance is also found to be enhanced. Here, as well, one has to put some additional effort, only during network training in terms of computations, but later on the system can be realized by inexpensive CMOS circuitry.

Chapter 9

Conclusions and Recommendations

9.1 Conclusions

In this thesis work, the problem of detection and classification of helicopters was studied. Simulated acoustic signals of different helicopters were utilized for identification and classification. Received acoustic signal was represented in the form of a parametric model. Different encoding schemes such as LPC coefficients, reflection coefficients, LSPs and cepstrum coefficients were applied to construct the feature vectors.

Helicopter sound, resulting due to the rotation of main and tail rotor blades, vortex noise, blade slap, engine vibration, Doppler's effect and atmospheric attenuation, is a complex phenomenon. We have investigated the application of ANN

as classifier to capture these complex attributes of the acoustic signal. Neural networks of varying sizes were trained using different encoding techniques. The classification system consists of two stages. At the first stage, a detection system decides about the presence of helicopter and at the second stage the helicopter type is determined.

At the detection stage ANN classifier was trained using the signal encoding parameters. The performance of different encoding schemes was compared with each other. The use of reflection coefficients as feature vectors has produced best results in terms of convergence rate and detection. It is also shown that the wavelet's de-noising capability and use of noisy signal along the clean signal in training stage has improved the performance. The performance results with these classifiers have shown better robustness in noisy conditions.

At the classification stage, ANN recognition systems with two hidden layers of neurons were trained. In one study four helicopters were used. Their acoustic signals were compressed in the form of feature vector using the encoding techniques mentioned above. Performance of different schemes was compared. The reflection coefficients based classifier was again the best in terms of classifying ability, network size and training time. It was also found the most robust against the background noise. A classifier was also trained using de-noised signal, obtained using wavelet transform. This results in improved performance.

In one study, two ANN classifiers were trained, one on the clean signal only and other on clean as well as noisy data to classify eight helicopters. Training on

noisy signals has added attributes of shielding the noise effects. This additive information learnt by the classifier has improved robustness in noisy conditions.

9.2 Further Research

Some recommendations for future work are as follow

- detection scheme can be investigated using real helicopter sounds at training and test stages
- fundamental blade passing frequency can be added to already discussed features to give additional attribute to the classifier
- work can be extended to more helicopters
- the detection and identification system can by implemented on low cost CMOS microcontrollers for remote operation
- system can be trained to track multi-targets
- cepsturm derivatives can be investigated to improve the performance
- use of cyclostationarity and time-frequency analysis can be investigated
- wavelet's compressed features can be used as input to the classifier

Appendix A

Assessing Stationarity of Signals

When the mean and the autocorrelation sequence of a random process do not change with time then the random process is called wide sense stationary or weakly stationary. The mean and autocorrlation sequence are defined as

$$m_x(t_1) = E[x(t_1)] \tag{A.1}$$

$$\varphi_x(t_1, t_2) = E[x(t_1)x(t_2)] \tag{A.2}$$

In our case we have divided the available helicopter acoustic signal into frames of duration 0.05, 0.1, 0.5 and 1 second. We have calculated mean and autocorrelation sequences for each frame. The computed autocorrelation sequences of each frame are plotted in (Fig.A1-A16). We have noticed that up to 0.5 second frame length, the decay of autocorrelation sequences has followed almost the same trend for successive frames. In case of a frame of duration 1 second, there is a distinct variation in the autocorrelation sequence's decay. This shows that helicopter's acoustic signal is non-stationary for 1 second frame length. In our application we have used 0.1 second frame due to limited availability of training data and subsequent computational requirements for feature extraction.

Frame length = 0.05 second

Mean of each frame is: -1.414, 1.173, -1.12, 0.218

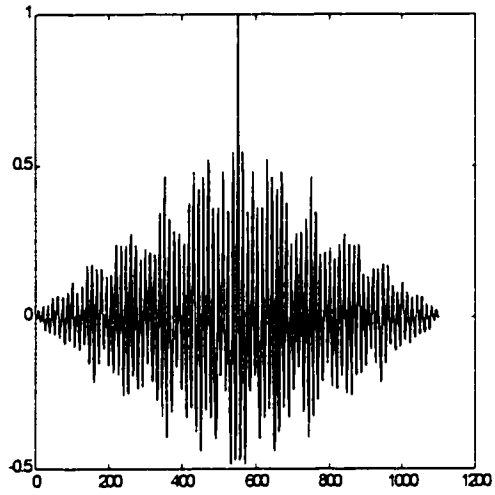


Fig.(A1)

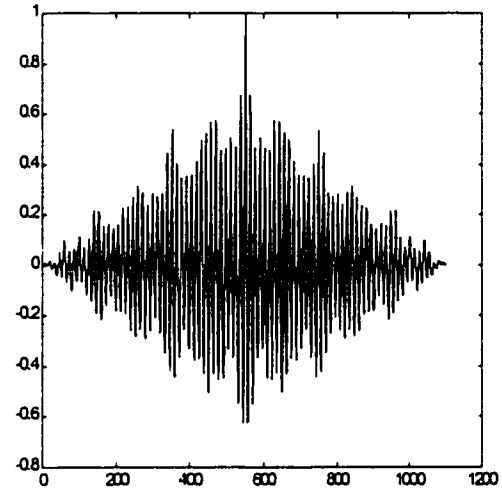


Fig.(A2)

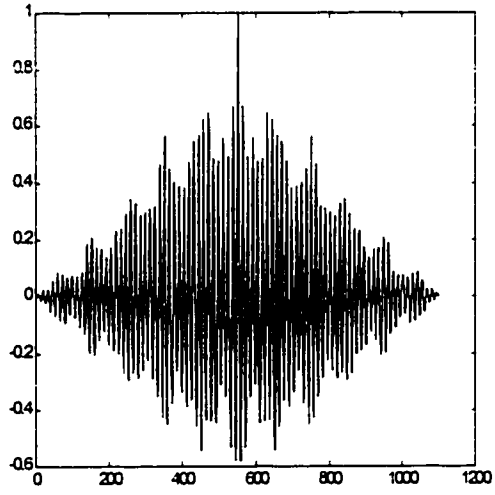


Fig.(A3)

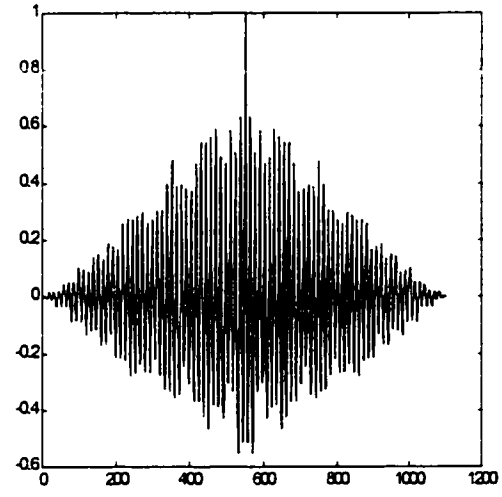


Fig.(A4)

Frame length = 0.1 second

Mean of each frame is: -12.06, -30.60, -30.75, 28.92

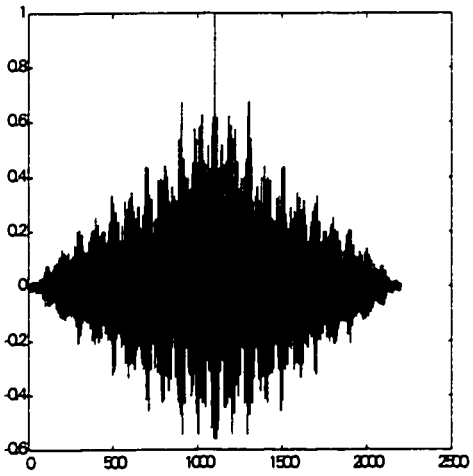


Fig.(A5)

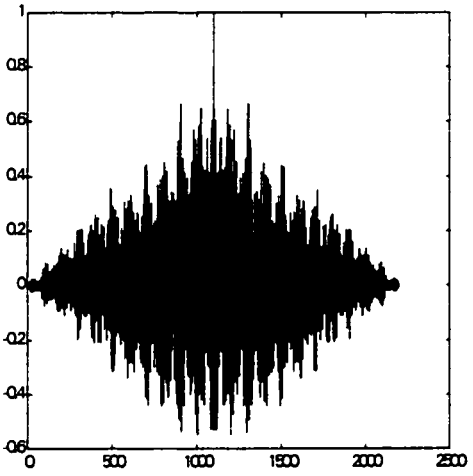


Fig.(A6)

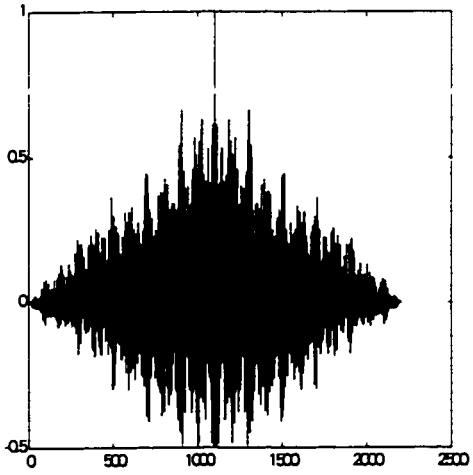


Fig.(A7)

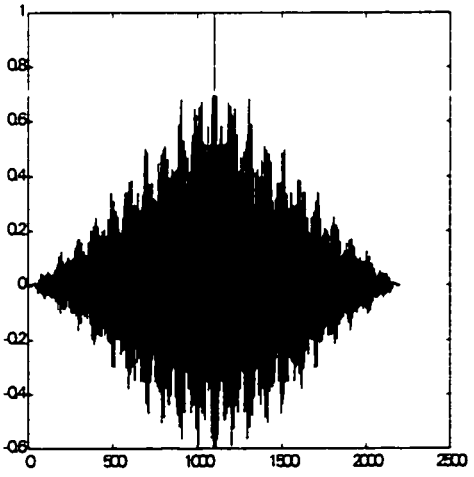


Fig.(A8)

Frame length = 0.5 second

Mean of each frame is: 6.52, 0.24, 3.37, 0.37

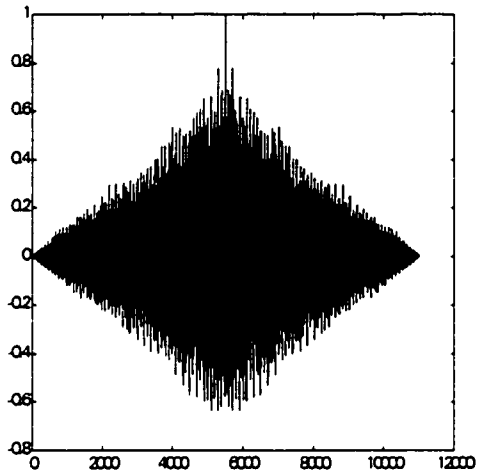


Fig.(A9)

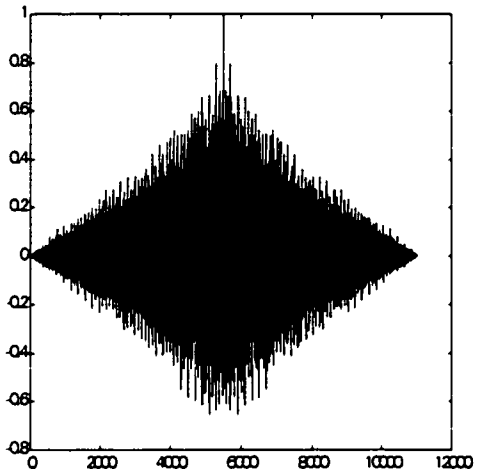


Fig.(A10)

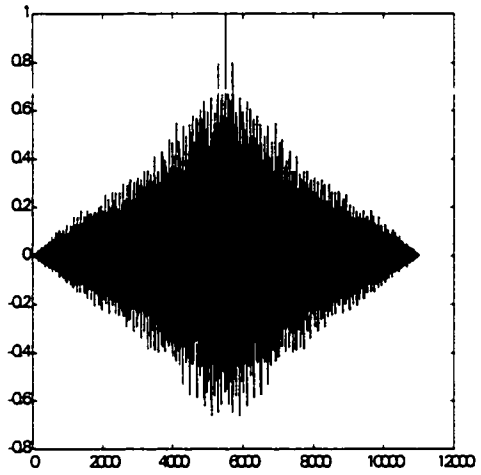


Fig.(A11)

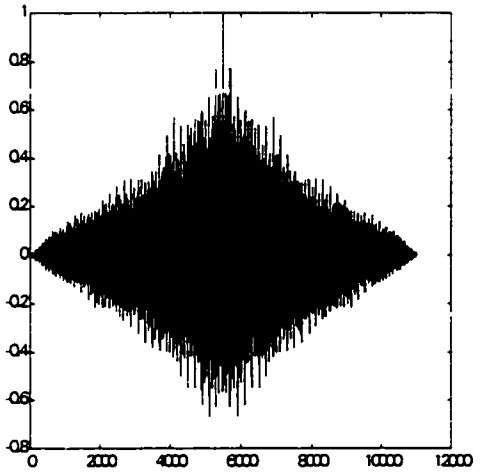


Fig.(A12)

Frame length = 1 second

Mean of each frame is: -5.99, 1.01, 1.97, 6.21

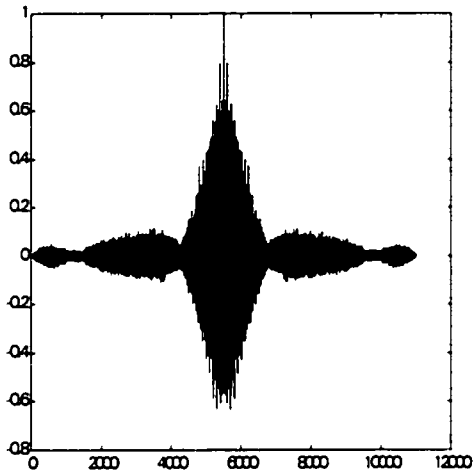


Fig.(A13)

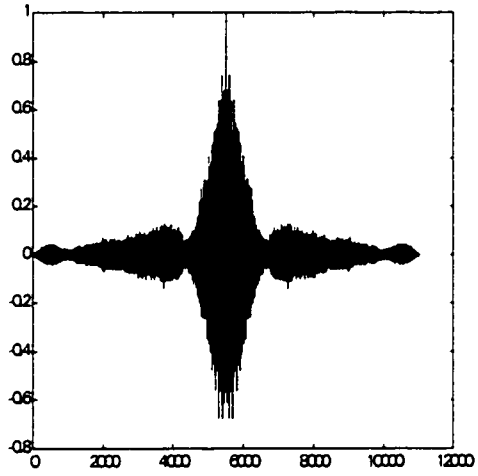


Fig.(A14)

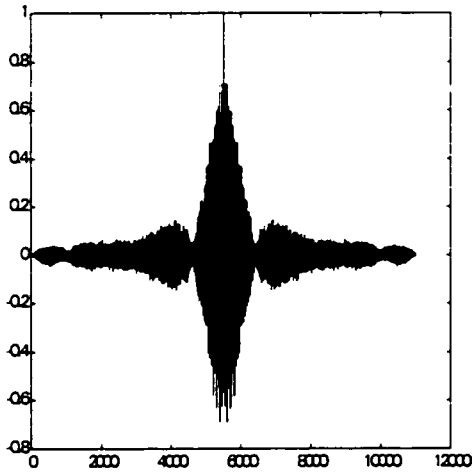


Fig.(A15)

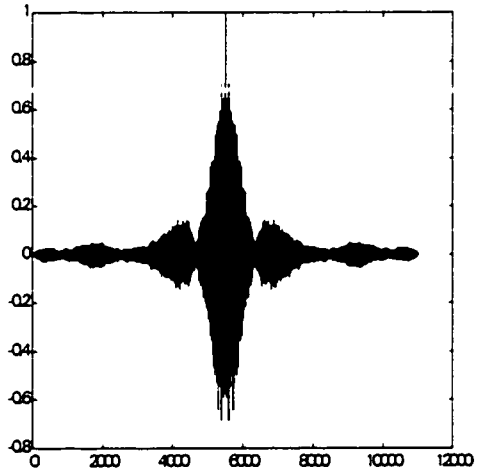


Fig.(A16)

Parametric Stationarity Tests

In many measurements, if the more complex properties (such as autocorrelation function or probability density function) of a signal change, so do some simple properties (such as the mean or the covariance). Thus, testing whether there is a change in these moments over time is an adequate method for assessing wide-sense stationarity. The most direct procedure is to divide a signal's record into multiple segments and to compare the means and variances of each segment among each other. Standard statistical tests are available for testing the equality of sample means and variances.

Equality of Means

Student's t-test is used to compare the means. The test statistics is a t statistic that uses a normalized value of the difference of the sample means and variances for populations 1 and 2. Let the sample means and variances of populations 1 and 2 be represented by the symbols $\hat{m}_1$, $\hat{\sigma}_1^2$, $\hat{m}_2$ and $\hat{\sigma}_2^2$. The test statistic is

$$T = \frac{\hat{m}_1 - \hat{m}_2}{\left[\left(\frac{1}{N_1} + \frac{1}{N_2}\right)\left(\frac{N_1\hat{\sigma}_1^2 + N_2\hat{\sigma}_2^2}{N_1 + N_2 - 2}\right)\right]^{\frac{1}{2}}} \tag{A.3}$$

T has a student's t-distribution with $\nu = N_1 + N_2 - 2$ degrees of freedom. The $(1-\alpha)$ confidence interval is

$$t_{\nu,1-\alpha/2} \le T \le t_{\nu,\alpha/2}$$

which is verified consulting t-distribution table.

Equality of Variances

The second parametric test for assessing the stationarity of signals requires testing the equality of variances of two sets of measurements. The test statistic F is the ratio of the two unbiased estimates of the variances,

$$F = \frac{\left(N_1 / \nu_1\right) \hat{\sigma}_1^2}{\left(N_2 / \nu_2\right) \hat{\sigma}_2^2} \tag{A.4}$$

F has two degrees of freedom, $\nu_1 = N_1 - 1$ and $\nu_2 = N_2 - 1$. The $(1-\alpha)$ confidence interval is

$$F_{\nu_1,\nu_2;1-\alpha/2} \le F \le F_{\nu_1,\nu_2;\alpha/2}$$

which is verified consulting F-distribution tables.

Appendix B

Levinson-Durbin Recursion

L-D recursion is a recursive-in-model-order solution for the autocorrelation equations. The vector matrix form of the autocorrelation equation is [22]

$$-\mathbf{R}\hat{\mathbf{a}}^{M} + \mathbf{r} = 0 \tag{B.1}$$

In the above expression superscript M determines the order of the solution, which needs proper tracking due to recursion.

Let us denote the prediction residual sequence due to the autocorrelation estimate $\hat{a}^{M}$ by $\varepsilon(n)$, which will result in

$$r(0) - \sum_{i=1}^{M} \hat{a}^{M}(i)\, r(i) = \frac{1}{N} \sum_{n=-\infty}^{\infty} \left[\varepsilon(n)\right]^{2} \tag{B.2}$$

Term on the right side is the scaled version of total energy in the prediction residual sequence. Rewrite this quantity ξ^{M} then Eq.(B.2) can be rewritten as

$$r(0) - \mathbf{r}^{T}\hat{\mathbf{a}}^{M} = \xi^{M} \tag{B.3}$$

Matrix form of Eq.(B.1) is

$$\begin{bmatrix} \mathbf{r} & \mathbf{R} \end{bmatrix} \begin{bmatrix} 1 \\ -\hat{\mathbf{a}}^{M} \end{bmatrix} = \mathbf{0} \tag{B.4}$$

Stack Eq.(B.3) on top of Eq.(B.4) to get

$$\begin{bmatrix} r(0) & \mathbf{r}^T \\ \mathbf{r} & \mathbf{R} \end{bmatrix} \begin{bmatrix} 1 \\ -\hat{\mathbf{a}}^M \end{bmatrix} = \begin{bmatrix} \xi^M \\ \mathbf{0} \end{bmatrix} \tag{B.5}$$

Let denote the augmented matrix on the left as $\widetilde{R}^M$. Above system of equations form the starting point of L-D recursion. Here the recursion is carried out for M=3.

The augmented matrix is

$$\begin{bmatrix} r(0) & r(1) & r(2) & r(3) \\ r(1) & r(0) & r(1) & r(2) \\ r(2) & r(1) & r(0) & r(1) \\ r(3) & r(2) & r(1) & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ -\hat{a}^3(1) \\ -\hat{a}^3(2) \\ -\hat{a}^3(3) \end{bmatrix} = \begin{bmatrix} \xi^3 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{B.6}$$

The augmented matrix, $\widetilde{R}^3$, is Toeplitz and the augmented matrix associated with second order recursion, $\widetilde{R}^2$, is embedded in it.

Let us assume that the augmented order-three solution vector in terms of order-two solution vector is as follows

$$\begin{bmatrix} 1 \\ -\hat{a}^3(1) \\ -\hat{a}^3(2) \\ -\hat{a}^3(3) \end{bmatrix} = \begin{bmatrix} 1 \\ -\hat{a}^2(1) \\ -\hat{a}^2(2) \\ 0 \end{bmatrix} - k(3) \begin{bmatrix} 0 \\ -\hat{a}^2(2) \\ -\hat{a}^2(1) \\ 1 \end{bmatrix} \tag{B.7}$$

The term $k(3)$ is called reflection coefficient. Putting this result in left-hand side of Eq.(B.6) remembering the embedded $\widetilde{R}^2$ we have

$$\begin{bmatrix} r(0) & r(1) & r(2) & r(3) \\ r(1) & r(0) & r(1) & r(2) \\ r(2) & r(1) & r(0) & r(1) \\ r(3) & r(2) & r(1) & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ -\hat{a}^3(1) \\ -\hat{a}^3(2) \\ -\hat{a}^3(3) \end{bmatrix} = \begin{bmatrix} \xi^2 \\ 0 \\ 0 \\ q \end{bmatrix} - k(3) \begin{bmatrix} q \\ 0 \\ 0 \\ \xi^2 \end{bmatrix} = \begin{bmatrix} \xi^3 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{B.8}$$

where $$q = r(3) - \sum_{i=1}^{2} \hat{a}^2(i) r(3-i) \tag{B.9}$$

Eq.(B.8) gives two scalar equations

$$\xi^2 - k(3)q = \xi^3 \tag{B.10}$$

$$q - k(3)\xi^2 = 0 \qquad \Rightarrow k(3) = \frac{q}{\xi^2} \tag{B.11}$$

Using Eq.(B.11), q can be eliminated from Eq.(B.10) and we get

$$\xi^3 = \xi^2\left\{1 - \left[k(3)\right]^2\right\} \tag{B.12}$$

Substituting Eq.(B.9) in Eq.(B.11) will result

$$k(3) = \frac{1}{\xi^2}\left\{r(3) - \sum_{i-1}^{2} \hat{a}^2(i)r(3-i)\right\} \tag{B.13}$$

The set of LP parameters for order-three is

$$\hat{a}^3(3) = k(3) \tag{B.14}$$

$$\hat{a}^3(i) = \hat{a}^2(i) - k(3)\,\hat{a}^2(3-i) \tag{B.15}$$

The recursion in short is given as

Initialization

ξ^0 = scaled total energy in the "error" from an "order 0" predictor = $r(0)$

Recursion for l=1,2,...,M

1. Compute the l^{th} reflection coefficient using

$$k(l) = \frac{1}{\xi^{l-1}}\left\{r(l) - \sum_{i-1}^{l-1} \hat{a}^{l-1}(i)r(3-i)\right\}$$

2. Generate the order-l set of LP parameters using

$$\hat{a}^l(l) = k(l)$$

$$\hat{a}^{l}(i) = \hat{a}^{l-1}(i) - k(l)\,\hat{a}^{l-1}(l-i)$$

3. Calculate the energy associated with order-l solution

$$\xi^{l} = \xi^{l-1}\left\{1 - \left[k(l)\right]^{2}\right\}$$

4. Return to step 1 with l replaced by l+1 if l<M

Appendix C

Computation of Reflection Coefficients

The steps involved in determining the predictor coefficients (PARCOR coefficients) using Burg's [10] algorithm are as follows:

1. Initially set $e^{(0)}(m) = s(m) = b^{(0)}(m)$
2. Determine $k_1 = a_1$ using the equation below

$$k_i = \frac{\sum_{n=0}^{N-1} e^{(i-1)}(n)\, b^{(i-1)}(n-1)}{\left\{\sum_{n=0}^{N-1}\left(e^{(i-1)}(n)\right)^2 \sum_{n=0}^{N-1}\left(b^{(i-1)}(n-1)\right)^2\right\}^{\frac{1}{2}}} \tag{C.1}$$

3. Determine forward and backward prediction errors $e^{(1)}(m)$ and $b^{(1)}(m)$ from the following equations respectively.

$$e^{(i)}(n) = e^{(i-1)}(n) - k_i\, b^{(i-1)}(n-1) \tag{C.2}$$

$$b^{(i)}(n) = b^{(i-1)}(n-1) - k_i\, e^{(i-1)}(n) \tag{C.3}$$

4. Set i=2
5. Determine $k_i = a_i^i$ from Eq.(C.1)
6. Determine $a_j^{(i)}$ for j=1,2,....,i-1 from the following Eq.

$$a_k^{(i)} = a_{k-1}^{(i-1)} - k_i\, a_{i-k}^{(i-1)} \tag{C.4}$$

7. Determine $e^{(i)}(m)$ and $b^{(i)}(m)$ from (C.2) and (C.3)

8. Set i=i+1

9. If i is less than or equal to p, go to step 5.

10. Procedure is terminated.

Appendix D

Computation of Line Spectral Pairs

The following computationally efficient algorithm can be employed for the calculation of Line Spectral Pair

1. The coefficients $e^{(i)}(m)$ and $b^{(i)}(m)$ for a 10th order LPC filter are given as:

$$\begin{aligned} p(i+1) &= a_{i+1} + a_{10-i} - p(i) \\ q(i+1) &= a_{i+1} - a_{10-i} + q(i) \end{aligned} \tag{D.1}$$

2. For root finding Newton's formula is used

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{D.2}$$

3. Same procedure is followed for finding the roots of two polynomials P and Q. Polynomial is treated as Z, with α_i, i=1,2,...N are the roots. Computation is done for a 10th order LPC polynomial i.e. N=5.

4. The polynomial $P_c(\omega)$ is referred as Z_n, with the n representing the order. With $x = \cos(\omega)$ the polynomial is:

$$\begin{aligned} Z_5 = \; & 16\,p(0)\,x^5 + 8\,p(1)\,x^4 + \left[-20\,p(0) + 4\,p(2)\right]x^3 + \left[-8\,p(1) + 2\,p(3)\right]x^2 \\ & + \left[5\,p(0) - 3\,p(2) + p(4)\right]x + p(1) - p(3) + 0.5\,p(5) \end{aligned} \tag{D.3}$$

The coefficients can be written with subscript representing the polynomial order as below

$$Z_5 = x^5 + z_5(1)\,x^4 + z_5(2)\,x^3 + z_5(3)\,x^2 + z_5(4)\,x \;+ z_5(5) \tag{D.4}$$

where $z_5(1) = \dfrac{8p(1)}{16p(0)}$, $z_5(2) = \dfrac{\{-20p(0) + 4p(2)\}}{16p(0)}$ and so on.

After getting the first root this root i.e., α_1, is factored out of the 5th order polynomial leaving a 4th order polynomial given by

$$Z_4 = x^4 + z_4(1)\,x^3 + z_4(2)\,x^2 + z_4(3)\,x \;+ z_4(4) \tag{D.5}$$

where $z_4(4) = \dfrac{-z_5(5)}{\alpha_1}$, *and* $z_4(j) = \dfrac{z_4(j+1) - z_5(j)}{\alpha_1}$ for j=1,2,...,4

Similarly the nth order polynomial and the coefficients after factoring out the kth root is given by:

$$Z_n = \sum_{i=0}^{n} z_n(i) x^{(n-i)} \tag{D.6}$$

where $z_n(0) = 1$, $z_n(n) = \dfrac{z_{n+1}(n+1)}{\alpha_k}$, *and* $z_n(j) = \dfrac{z_n(j+1) - z_{n-1}(j)}{\alpha_k}$ for j=n-1,...1 and n=N-1,...,1

1. We also need the values of the differential of the polynomial Z_n, which is represented by Z_n'. The differential of the 5th order polynomial is given by:

$$Z_5' = z_5'(1)\,x^4 + z_5'(2)\,x^3 + z_5'(3)\,x^2 + z_5'(4)\,x \;+ z_5'(5) \tag{D.7}$$

where $z_5'(1) = 5$, $z_5'(2) = 4z_5(1)$ and so on.

In general the nth order differential polynomial is given as

$$Z'_n = \sum_{i=1}^{n} z_n(i)\, x^{(n-1)} \tag{D.8}$$

where $z'_n(i) = (n - i + 1)\ z_n(i-1)$, for I=1,...,n and $z_n(0) = 1$

References

[1] Mammone R.J., Xiaoyu Zhang, Ramachandran R.P., "Robust speaker recognition-A feature based approach" *IEEE Signal Processing Magazine*, September 1996.

[2] Junqua J-C., Haton J-P., "Robustness in automatic speech recognition-Fundamentals and Applications" *Kluwer Academic Publishers, Boston*, 1996.

[3] Elshafei Ahmed, M.S. Ahmed, "Helicopter Recognition using Neural Computation" *National Computer Conference, Dhahran*, v1, p 121-133, 1997.

[4] Cabell R.H., Fuller C.R., O'Brien W.F., "Identification of helicopter noise using a neural network" *AIAA Journal*, v30, n3, p 624-630, March 1992.

[5] Cecil F. Hess, "Optical microphone for the detection of hidden helicopters" *AIAA Journal*, v30, n11, p 624-630, November 1992.

[6] Meir Feder, "Parameter estimation and extraction of helicopter signals observed with a wide-band interference" *IEEE Trans. on Signal Processing*, v41, n1, p 232-244, January 1993.

[7] Lin Zhiping, "Detection of Helicopter signals using cyclostationarity" *Proceedings ICASSP*, v3, p 1952-1955, 1995.

[8] Vezzosi M., Ehrmann G., "Helicopter detection by analysis of the acoustic noise emitted by the main rotor" *Journal De Physique*, v4, n5, pt2, p 1339-1342, May 1994.

[9] Cabell R.H., Fuller C.R., O'Brien W.F., "Neural Network for the identification of measured Helicopter noise" *Journal of American Helicopter Society*, v38, n3, p 64-72, July 1993.

[10] Hemdal J.F., "An investigation of acoustic and seismic sources, final report" *Willow Run Lab. CR 86, Ann Arbor, MI*, February 1973.

[11] Mori S., "Tracking and classifying multiple targets without a priori identification" *IEEE Trans. Automatic Control*, vAC-31, n5, p 401-409, May 1986.

[12] Lippman R., "An introduction to computing with neural nets" *IEEE ASSP Magazine*, p 4-22, April 1987.

[13] Leverton J.W., "The sound of rotorcraft" *RAeS Aeron. Journ.*, v75, p 385-387, 1971.

[14] Splettstoresser W.R., Schultz K.J., Boxwell D.A., Schmitz F.H., "Helicopter model rotor-blade vortex interaction noise: Stability and parametric variations" *10th European Rotorcraft Forum*. Paper 18, 1984.

[15] Michael J.T. Smith, "Aircraft Noise" *Cambridge University Press, Cambridge*, 1989.

[16] Wrigth S.E., "Sound radiation from a lifting rotor generated by asymmetric disk loading" *J. Sound Vib.*, v9, p 223-240,1969.

[17] Lighthill M.J., "On sound generated aerodynamically Part I General Theory" *Proc. Royal Soc. (London)*, A211, p 564-587.1952.

[18] Gutin L., "On the sound field of a rotating propeller" *NASA, TM 1195*, 1948.

[19] Lowson M.V., Ollerhead J.B., "A theoretical study of helicopter rotor noise" *J. Sound Vib.*, v9, p 197-222, 1969.

[20] Barry B., Moore C.J., "Subsonic fan noise" *J. Sound Vib.*, v17, p207-220, 1971.

[21] Richards E.J. Mead D.J., "Noise and acoustic fatigue in aeronautics" *John Wiley and Sons Ltd., London*, 1968.

[22] Morfey C.L., Tanna H.K., "Sound radiation from a point source in circular motion" *J. Sound Vib.*, v15, p 325-351, 1971.

[23] Glegg S., "Fan Noise" *In White, R.G. and Walker, J.G. (Edrs.), Noise and Vibration, Ellis Horwood Ltd.*, p 439-462.

[24] Simons I.A., "Oscillatory aerodynamic loads on helicopter rotor blades in the hover" *Univ. Southampton Intl. Rep. I.S.V.R.*, Feb. 1966, (Unpublished).

[25] Loewy R.G., Sulton L.R., "A theory of predicting the rotational noise of lifting rotor in forward flight including a comparison with experiment" *J. Sound Vib.*, v4, p 305, 1966.

[26] Davidson I.M., Hargest T.J., "Helicopter Noise" *J. Roy. Aeron. Soc.*, May 1965.

[27] Goddar J.O., Stuckey, T.J., "Investigation and prediction of helicopter rotor noise (Part I- Wessex Whirl Tower Results)" *J. Sound Vib.*, v5, p50, 1967.

[28] Bell Helicopter Co. "A study of the origin and means of reducing helicopter noise" *United States Tech. Rep. T.C.R.E.C.*, p 62-73, Nov. 1962.

[29] Hubbard H.H., Lassiter L.N., "Sound from a two blade propeller at supersonic tip speeds" *NASA Rep. 1079*, 1952.

[30] Beranek L.L. (Edited), "Noise and vibration control" Ch#7, *McGraw Hill Book Co.*, New York, 1971.

[31] Harris C.M., "Absorption of sound in air versus humidity and temperature" *J. Acoust. Soc. Amer.*, v40, p 148-159, 1966.

[32] Knudsen V.O., Wilson J.V., Anderson N.S., "The attenuation of audible sound in fog and smoke" *J. Acoust. Soc. Amer.*, v20, p 849-857, 1948.

[33] Kragh J., Jakobsen J., "Propagation of wind turbine noise, outline of a prediction method" *DELTA Acoustic and Vibration*, AV 320/95, p 1-30, 1995.

[34] Satio S., Itakura F., "The theoretical consideration of statistically optimum methods for speech spectral density" *Report No. 3107, Electrical Communication Laboratory, N.T.T., Tokyo*, 1966 (In Japans).

[35] Atal B.S., Schroeder M.R., "Adaptive predictive coding of speech signals" *The Bell System Technical Journal*, p 1974-1986, October 1970.

[36] Makhoul John, "Linear prediction: a tutorial review" *Proceedings of the IEEE*, v63, n4, p 561-580, April 1975.

[37] Steve Young, "A review of the large-vocabulary continuous-speech recognition" *IEEE Signal Processing Magazine*, September 1996.

[38] Deller J.R., Proakis J.G., Hansen J.H.L., "Discrete-Time Processing of Speech Signals" *Macmillan Publishing Company, New York*, 1993.

[39] Makhoul J., "Stable and efficient lattice method for linear prediction" *IEEE Trans. on ASSP*, v ASSP25, n 5, p 423-428, October 1977.

[40] Markel J.D., Gray A.H., "Linear Prediction of Speech" *Springer-Verlag Publishing Company, New York*, 1976.

[41] Rabiner L.R., Schafer R.W., "Digital processing of speech signal" *Prentice-Hall, Inc., Englewood Cliffs, New Jersey*, 1978.

[42] Burg J., "A new analysis technique for time series data" *Proc. NATO Advanced Study Institute on Signal Proc., Enschede Netherlands*, 1968.

[43] Sugamura N., Itakura F., "Speech analysis and synthesis methods developed at ECL in NTT -from LPC to LSP-" *Speech Communication*, v 5, p 199-215, 1986.

[44] Itakura F., "Line spectrum representation of linear predictor coefficients of speech signals" *J. Acoust. Soc. Amer.*, v57, s35(A), 1975.

[45] Peter Kabal, Ravi Prakash Ramachandran, "The computation of line spectral frequencies using Chebyshev polynomials" *IEEE Trans. on ASSP*, v ASSP34, n 6, p 1419-1425, December 1986.

[46] Frank K. Soong, Biing-Hwang Juang, "Line spectral pairs and speech data compression" *In Proc. Int. Conf. Acoust., Speech, Signal Processing, San Diego, CA*, p 1.10.1-1.10.4, March 1984.

[47] Atal B.S., "Effectiveness of linear predictive characteristics of the speech wave for automatic speaker identification and verification" *J. Acoust. Soc. Amer.*, v55, p 1304-1312, June 1974.

[48] Furi S., "Cepstral analysis technique for automatic speaker verification" *IEEE Trans. Acoust., Speech, Signal Processing*, v29, p 254-272, April 1981.

[49] Mammone R.J., Xiaoyu Zhang, Ramachandran R.P., "Robust speaker recognition-A feature based approach" *IEEE Signal Processing Magazine*, September 1996.

[50] Quen-Zang Wu, I-Chang Jou, Suh-Yin Lee, "On-line signature verification using LPC cepstrum and neural networks" *IEEE Trans. On Systems, Man and Cybernetics-Part B: Cybernetics*, v27, n1, p 148-153, Feb. 1997.

[51] d' Alessandro C., "Speech analysis and synthesis using an auditory-based wavelet representation" *In ETRW: Comparing Signal Representation, Sheffield, England*, p 31-38,1992.

[52] Ambikairajah E., Keane M., Kilmartim L., Tattersall G., "The application of the wavelet transform for speech processing" *In EUROSPEECH*, p 151-154, 1993.

[53] Wilde S., Curtis K., "The wavelet transform for speech analysis" *In ICSLP*, p 1621-1624, 1992.

[54] Teolis A., Benedetto J., "Noise suppression using a wavelet model" *In ICASSP*, p17-20, 1994.

[55] Park S.W., "Speech compression using ARMA model and wavelet transform" *In ICASSP*, p 209-212, 1994.

[56] Kadambe S., Boudreaux-Bartels G., "A comparison of wavelet functions for pitch detection of speech signals" *In ICASSP*, p 449-452, 1991.

[57] Hwang W.L., Mallat S., "Singularities and noise discrimination with wavelets" *In ICASSP*, p 377-380, 1992.

[58] Leung S., Wong O., Lai K., "Decomposition of the LPC excitation using wavelet functions" *In EUROSPEECH*, p 1327-1331, 1991.

[59] Favero R., Gurgen F., "Using wavelet dyadic grids and neural networks for speech recognition" *In ICSLP*, p 1539-1542, 1994.

[60] Daubechies I., "Ten lectures on wavelets" *CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics*, 1992.

[61] Misiti M., Misiti Y., Oppenheim G., Poggi J-M., "Wavelet Toolbox" *The Math Works, Inc., 24 Prime Park Way, Natick, Mass.*, 1996.

[62] Mallat S., "A theory for multiresolution signal decomposition: the wavelet representation" *IEEE Trans. On Pattern Analysis and Machine Intell.*, v11, v7, p 674-694, July 1989.

[63] Rosenfeld A. (Edited), "Multiresolution techniques in Computer vision" *Springer-Verlag, New York*, 1984.

[64] Strang G., Nguyen T., "Wavelet and filter bank" *Wellesley-Cambridge Press*, 1996.

[65] Vetterli M, Herley C., "Wavelets and filter banks: Theory and design" *IEEE Trans. On Signal Proc.*. 1992.

[66] McCulloch W.S., Pitts W., "A logical calculus of the ideas immanent in nervous system" *Bulletin of Mathematical Biophysics*, v5, p 115-133, 1943.

[67] Uttley A.M., "A theory of mechanism of learning based on the computation of conditional probabilities" *Proc. Of the 1st Int. Conf. On Cybernetics, Namur, Gauthier-Villars, Paris*, 1956.

[68] Rosenblatt F., "The perceptron: A probabilistic model for information storage and organization in brain" *Psychological Review*, v65, p 386-408, 1958.

[69] Minsky M.L., Papert S.A., "Perceptrons" *Cambridge, MA: MIT Press*, 1969.

[70] Werbos P.J., "Beyond regression: New tools for prediction and analysis in behavioral sciences" *Ph.D. Thesis, Harvard University, Cambridge, MA*, 1974.

[71] Hopfield J.J., "Neural networks and physical systems with emergent collective computational abilities" *Proc. Of the National Academy of Sciences of the U.S.A.*, v79, p 2554-2558, 1982.

[72] Rumelhart D.E., Hinton G.E., Williams R.J., "Learning internal representation by back Propagation" *In Parallel Distributed Processing: Exploration in the Microstructure of Cognition, V. I , Cambridge, MA: Bradford Books*, 1986.

[73] Simon Haykin, "Neural Networks; A Comprehensive Foundation" *Macmillan Publishing Company Englewood Cliffs, NJ*, 1994.

[74] Widrow B., Winter R.G., Baxter R.A., "Layered Neural Nets for Pattern Recognition." *IEEE Trans. Acoust. Speech and Signal Proc.*, V. 36, N 7, p 1109-1118, 1988.

[75] Widrow B., Winter R.G., Baxtar R.A., "Layered neural nets for pattern recognition" *IEEE Trans. ASSP Proc.*, v36, n7, p 1109-1118, 1988.

[76] Rippnann R.P., "An introduction to computing with Neural Nets" *IEEE Acoustic Speech and Signal Processing Magazine*, p 4-22, April 1987.

[77] Hornik K., Stinchcombe M., White H., "Multilayer feed-forward network are universal approximators" *Neural Networks*, v2, p 359-366, 1989.

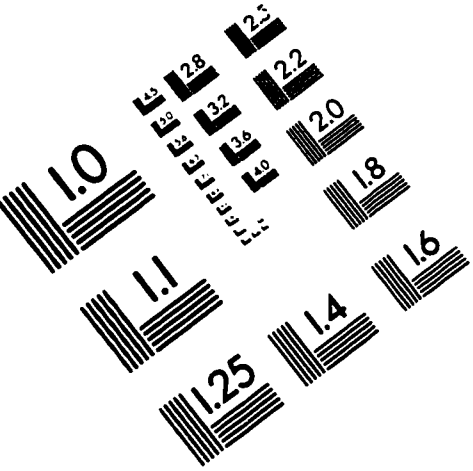
[78] Mansour D., Juang B.-H., "A family of distortion measure based upon projection operation for robust speech recognition" *IEEE Trans. ASSP*, v ASSP-37, n 11, p 1659-1671, 1989.

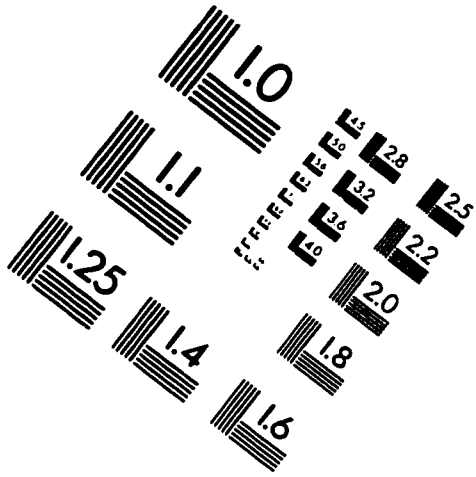
[79] Bateman D., Bye D., Hunt M., "Speech contrast normalization and other techniques for speech recognition in noise" *In ICASSP*, p I.241-I.244,1992.

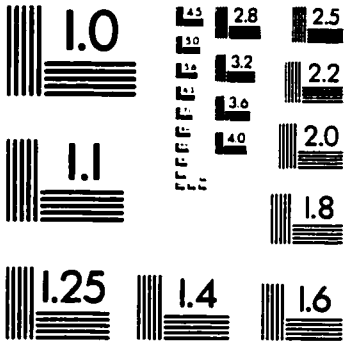
Vita

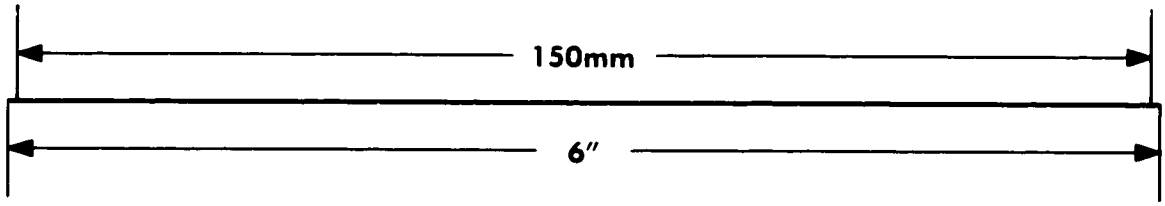
- Sohail Akhtar
- Born in Wah Cantt, Pakistan.
- Received Bachelor's degree in Electrical Engineering from the University of Engineering and Technology, Lahore, Pakistan in March, 1993.
- Worked in Wah Nobel Chemicals Pvt. Ltd. Wah Cantt, Pakistan, for three years.
- Completed Master's degree requirements at King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia in December, 1998.

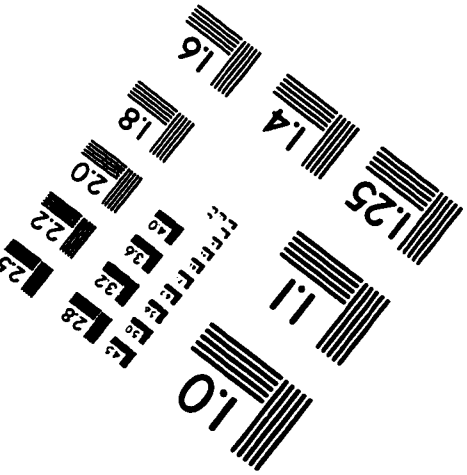
IMAGE EVALUATION
TEST TARGET (QA-3)

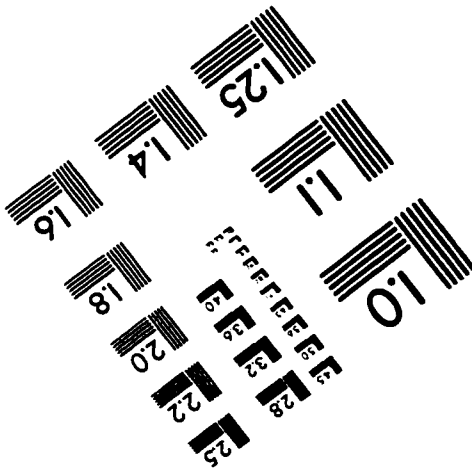












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989